作業1

Advanced Statistical Inference II

Homework 1: Review of probability language and some concepts in statistics.

Due Date: March 2nd, 2017

1. (Review operation of random variables) Let X and Y be independent Poisson random variables with mean μ and λ respectively. Find the distribution of $X+Y$ and conditional distribution of X given that $X+Y=n$.

2. (Review change of variable and integration) Let X and Y be independent exponential random variables with rate α. Find the densities of the random variables X^3, $|X-Y|$, and $\min(X, Y^3)$.

3. (Review college level probability and statistics) In R^2 let the two coordinate axes be denoted by the x-axis and the y-axis. An isoceles triangle is formed by a unit vector in the (positive) x-direction and another unit vector in a random direction. Find the distribution of the length of the third side.

4. (Key fact) Let F be a cumulative distribution function which is continuous and strictly increasing. Let $X \sim U[0,1]$ and define $Y = F^{-1}(X)$. Compute the distribution function of Y.

5. (Order statistics) Let X_i, $1 \le i \le 3$ be IID $U[0,1]$ and let the corresponding order statistics be $X_{(i)}$. Let $U_1 = X_{(1)}/X_{(2)}$ and $U_2 = X_{(2)}/X_{(3)}$. Show that U_1 and U_2 are independent. Your can refer to the note at

   `http://www.math.ntu.edu.tw/~hchen/teaching/LargeSample/notes/noteorder.pc`

6. (Order statistics) It is assumed that the lifetimes of electric bulbs have an exponential distribution with an unknown expectation α^{-1}. To estimate α, a sample of n bulbs is taken but one only observes the lifetimes of these bulbs $X_{(1)} < X_{(2)} < \cdots < X_{(r)}$. Let

$$U = \frac{n-r}{r} X_{(r)} + \frac{1}{r}\sum_{i=1}^{r} X_{(i)}.$$

   Show that $E(U) = \alpha^{-1}$ and $Var(U) = \alpha^{-2}/r$.

7. Let X_1 and X_2 be independent and identically distributed. Show that

$$P(|X_1 - X_2| > t) \le 2P(|X_1| > t/2).$$

8. (Review MLE and asymptotic analysis) Let $X_1, \ldots, X_n$ be i.i.d. random variables with common density function

$$f_\theta(x) = \frac{\theta}{(1+x)^{\theta+1}}, \quad x > 0, \theta > 0.$$

   (a) Find the maximum likelihood estimator of θ, denoted as $\hat{\theta}_n$.

   (b) Find the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$.

   (c) Find a function g such that, regardless the value of θ, $\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \to N(0,1)$ in distribution.

定 $Y_1 \sim Y_{n+1} \overset{iid}{\sim} \exp(1)$

$X_1 = \frac{Y_1}{Y_1+\cdots+Y_{n+1}}$, $X_2 = \frac{Y_1+Y_2}{Y_1+\cdots+Y_{n+1}}$, $\ldots$ $X_n = \frac{Y_1+\cdots+Y_n}{Y_1+\cdots+Y_{n+1}}$

$(X_1, \ldots X_n)$ vs $(U_{(1)}, U_{(2)} \ldots U_{(n)})$ 同分佈

証. $X_{n+1} = Y_1+\cdots+Y_{n+1}$, 變數變換

9. (Law of large numbers and importance sampling) The following algorithm has been proposed to find $P(Z > 5) = 1 - \Phi(5)$ where Z is a standard normal random variable.

Step 1. Sample $x_1, x_2, \ldots, x_m$ from $N(5, 1)$.

Step 2. Calculate

$$\hat{p}_{IS} = \frac{1}{m}\sum_{i=1}^{m} I(x_i \geq 5)\frac{\phi(x_i)}{\phi(x_i - 5)}$$

where $\phi(\cdot)$ is the density function of Z.

(a) Is $\hat{p}_{IS}$ an unbiased estimator of $1 - \Phi(5)$? Justify your answer.

(b) Determine the variance of $\hat{p}_{IS}$.

10. (Over Dispersion and Mixture Distribution) The Conway-Maxwell-Poisson distribution has the probability function

$$P(Y = y) = \frac{\lambda^y}{(y!)^\nu}\frac{1}{Z(\lambda, \nu)}, \quad y = 0, 1, 2, \cdots$$

where

$$Z(\lambda, \nu) = \sum_{i=1}^{\infty}\frac{\lambda^i}{(i!)^\nu}.$$

(a) Place this distribution in an exponential family form with respect to both parameters, and identify all the relevant components.

(b) Determine ν such that the Conway-Maxwell-Poisson distribution is reduced to the Poisson distribution.

(c) Explain why this distribution can be used to model overdispersion for count data.

To have a better understanding on the term *overdispersion*, please refer to

`http://data.princeton.edu/wws509/notes/c4a.pdf`

11. Long (1990, 1997) gave data on the number of publications by 915 doctoral candidates in biochemistry. The numbers of published papers during last three years of Ph.D. are $0, 1, 2, 3, \cdots, 12$, 16 and 19. The corresponding number of students from 275, 246, 178, 84, 67, 27, 17, 12, 1, 2, 1, 1, 2, 1, 1. When the number of publications can be modelled as a Poisson random variable with mean λ, can you be certain that they come from Poisson based on the fact that $E(X) = Var(X) = \lambda$.

12. The double-exponential distribution has pdf $f(x) = (\lambda/2)\exp(-\lambda|x|$, for fixed λ ¿ 0.

(a) Find the MGF $M_X(t)$ of a double exponential. For which t is it finite?

(b) Let U and V are independent and identically distributed random variables with exponential distribution with mean 1,and 0find the MGF $M_Y(t)$ of $Y = U - V$. What is the distribution of Y?

(c) Find the mean and variance of a double exponential. If $X_1, \cdots, X_n$ are i.i.d. double exponentials, find the MGF $M_n(t)$ of the standardized mean,

$$W_n = \frac{\bar{X} - E[X_1]}{\sqrt{Var(X)/n}}.$$

(d) What is the limit of $M_n(t)$ as $n \to \infty$? What distribution has this function for its MGF?

13. Let $X_1, \cdots, X_n$ be IID normal random variables with mean μ and variance 1. Here $\mu \geq 1$.

(a) Find the MLE of μ which is denoted by $\hat{\mu}_{mle}$.

(b) When the true $\mu > 0$, determine the distribution of $\sqrt{n}(\hat{\mu}_{mle} - \mu)$ when n is large.

$\mu > 1$

(c) When the true $\mu = 0$, determine the distribution of $\sqrt{n}(\hat{\mu}_{mle} - \mu)$ when n is large.

$\mu = 1$

14. Suppose U and V are independent with exponential distribution with parameter λ. (A random variable T is exponentially distributed with parameter λ if its density is given by $f(t) = \lambda \exp(-\lambda t)$ with support $T > 0$.) Define $X = U + V$ and $Y = UV$.

(a) Derive the joint density of (X, Y).

(b) Find the best linear predictor of Y given X.

(c) Find the best predictor of Y given X.

15. Suppose that X and Y have a joint pdf given by

$$f_{X,Y}(x, y) = \begin{cases} 2 & 0 < x < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $E[Y|X = x_0]$.

(b) Plot the CEF (conditional expectation function) $E[Y|X = x]$. Can you explain heuristically why the function has this particular form?

(c) Find $E[YX^3 + 1|X = x_0]$.

(d) Find $V(Y|X = x_0)$. Does its dependence on x_0 make sense to you?

16. The Weibull cumulative distribution function is

$$F(x) = 1 - \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right], \quad x \geq 0, \alpha > 0, \beta > 0.$$

(a) Find the density function.

(b) Show that if W follows a Weibull distribution, then $X = (W/\alpha)^{\beta}$ follows an exponential distribution.

(c) How could Weibull random variables be generated from a uniform random number generator?

17. Let $X_1, \cancel{X_2} X_2 \sim Uniform(0, \theta)$ where $\theta > 0$.

(a) Find the distribution of (X_1, X_2) given T where $T = \max\{X_1, X_2\}$.

(b) Show that $X_1 + X_2$ is not sufficient.

18. Let $X_1, \cdots, X_n \sim Uniform(-\theta, 2\theta)$ where $\theta > 0$. Find the likelihood function.

19. Consider four observations -1, 0, 0.5, and 3 and evaluation at $x = 0$, $x = 0.5$, and $x = 1$. Using a bandwidth of 1, determine Gaussian kernel density estimate at $x = 0$, $x = 0.5$, and $x = 1$. Note that the resulting estimate at $x = 0$ should be 0.249.

20. You are given a kernel $K(\cdot)$ which satisfies $K(u) \geq 0$, $\int K(u)du = 1$, $\int uK(u)du = 0$, $\int u^2K(u)du = \sigma_K^2 < \infty$. You are also given a bandwidth $h > 0$, and a collection of n univariate observations $x_1, x_2, \cdots, x_n$. Assume that the data are independent samples from some unknown density f.

   (a) Give the formula for $\hat{f}_h$, the kernel density estimate corresponding to these data, this bandwidth, and this kernel.

   (b) Find the expectation of a random variable whose density is $\hat{f}$, in terms of the sample moments, h, and the properties of the kernel function.

   (c) Find the variance of a random variable whose density is $\hat{f}_h$, in terms of the sample moments, h, and the properties of h the kernel function.

   (d) How must h change as n grows to ensure that the expectation and variance of $\hat{f}_h$ will converge on the expectation and variance of f?

21. Let $X_1, \cdots, X_n$ be a random sample from the density

$$f(x|\theta) = \theta x^{\theta-1} I_{(0,1)}(x).$$

   The parameter space is $\Theta = (0, \infty)$.

   (a) Verify that $-\log X_1 = Y$ has an exponential distribution.

   (b) Find the Cramer Rao lower bound for unbiased estimators of $\tau(\theta) = 1/\theta$.

   (c) Show that $-\sum_{i=1}^n \log X_i/n$ is an UMVUE of $1/\theta$.

22. Let $X_1, \cdots, X_n$ be iid $U[0, \theta]$, and suppose that we want to estimate θ.

   (a) Show that $X_{(n)} = \max_{1 \leq i \leq n} X_i$ is sufficient for θ.

   (b) Let $\tilde{\theta} = 2X_1$, show that $\tilde{\theta}$ is an unbiased estimator for θ.

   (c) Find $E(\tilde{\theta}|X_{(n)})$ and show that it is a UNVUE of θ

23. Suppose $X_1, \cdots, X_n$ are iid Poisson(λ), and let $\theta = \exp(-\lambda)$ which is $P(X_1 = 0)$.

   (a) Show that $T = \sum_{i=1}^n X_i$ is sufficient for θ.

   (b) Consider an estimator $\tilde{\theta} = 1_{X_1=0}$. Show that $\tilde{\theta}$ is an unbiased estimator of θ.

   (c) Show that $E(\tilde{\theta}|T = t) = (1 - 1/n)^{\sum_i X_i}$.

24. Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a sample from an exponential distribution with individual densities

$$f(x; \theta) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right), \quad x > 0,$$

   where $\theta > 0$ is unknown.

θ^2 ?

(a) Show that $\tilde{\theta} = X_1^2/2$ is an unbiased estimator of $g(\theta) = \theta$;

(b) Show that $t(\mathbf{X}) = \sum_{i=1}^n X_i$ is sufficient for θ.

(c) Rao-Blackwellize $\tilde{\theta}$ to find an improved unbiased estimator of $g(\theta)$ which is denoted by $\hat{\theta}_u$;
Hint: You may use without proof that the distribution of $U = X_1/(\sum_{i=1}^n X_i)$ follows a Beta-distribution $\mathcal{B}(1, n-1)$, where $\mathcal{B}(\alpha, \beta)$ is the distribution with density

$$f(t; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1}(1-t)^{\beta-1}, \quad 0 < t < 1.$$

(d) Find the MLE of θ^2 and denote it by $\hat{\theta}_{mle}$. Compare the mean square error of the two estimates.

25. Suppose $X \sim N(0, \sigma^2)$. Note that the density function of X can be parametrized by a single parameter σ and

$$f(x|\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right) I(x \in R).$$

(a) Specify $\eta(\sigma)$, $T(x)$, $\psi(\sigma) = \log \sigma$, $h(x) = (1/\sqrt{2\pi})I(x \in R)$ to represent the density in the form

$$f(x|\sigma) = \exp(\eta(\sigma)T(x) - \psi(\sigma))h(x),$$

for any $\sigma \in R^+$.

(b) Suppose that we have an iid sample $X_1, X_2, \cdots, X_n \sim N(0, \sigma^2)$. Then write the joint density of $X_1, X_2, \cdots, X_n$ in the same general form

$$f(x_1, x_2, \cdots, x_n|\sigma) = \exp\left(\eta(\sigma)T(x_1, x_2, \cdots, x_n) - \psi(\sigma)\right) h(x_1, x_2, \cdots, x_n).$$

Show that the sufficient statistic is $\sum_{i=1}^n X_i^2$.

(c) Using the fact that the sum of squares of n independent standard normal variables is a chi square variable with n degrees of freedom, we have that the density of $T(X_1, X_2, \cdots, X_n)$ is

$$f_T(t|\sigma) = \frac{\exp(-t/(2\sigma^2))t^{n/2-1}}{\sigma^n 2^{n/2}\Gamma(n/2)} I(t > 0).$$

Please identify $\eta(\sigma)$, $S(t)$, $\psi(\sigma)$ and $h(t)$.

$\to$?

作業2

Advanced Statistical Inference II

Homework 2: Review of probability language, asymptotic analysis and some concepts in statistics.

Due Date: March 20th, 2017

1. Suppose X is a $Binomial(n, p)$ random variable, $0 < p < 1$.

   (a) Find the mle for $p(1-p)$.

   (b) Show that the mle is not unbiased for $p(1-p)$.

   (c) Construct an unbiased estimator for $p(1-p)$ using this mle given in (a).

2. (complete) Suppose X is a $Binomial(2, p)$ random variable where p can take only two values $1/2$ and $1/4$. Show that this family is not complete by constructing a non-zero function $g(X)$ whose expectation is zero for both $p = 1/2$ and $1/4$.

3. (order statistic) Let X_1, X_2, X_3 be a random sample of size three from the uniform distribution $U(0, \theta)$, where $\theta > 0$ is an unknown parameter. Let $X_{(1)}, X_{(2)}, X_{(3)}$ be the corresponding order statistics.

   (a) Find the marginal pdf $X_{(1)}$ and show that $X_{(1)}/\theta$ is distributed as $Beta(1, 3)$.

   (b) Compute $E[X_{(1)}]$. Construct an unbiased estimator for θ using $X_{(1)}$.

   (c) Show that $X_{(3)}/X_{(1)}$ is independent of $X_{(3)}$.

4. (order restricted inference) Let X be a single observation from the $N(\theta, 1)$ distribution, $\theta > 0$. Notice that the unknown mean is assumed to be positive.

   (a) Obtain the maximum likelihood estimator (MLE) of θ, $\hat{\theta}$. (Be sure to consider the parameter space.)

   (b) Show that $E(\hat{\theta}) \neq \theta$.

5. (Review change of variable and integration) Let X and Y be independent exponential random variables with rate α. Find the densities of the random variables X^3, $|X-Y|$, and $\min(X, Y^3)$.

6. (Order statistics) It is assumed that the lifetimes of electric bulbs have an exponential distribution with an unknown expectation α^{-1}. To estimate α, a sample of n bulbs is taken but one only observes the lifetimes of these bulbs $X_{(1)} < X_{(2)} < \cdots < X_{(r)}$. Let

$$U = \frac{n-r}{r} X_{(r)} + \frac{1}{r} \sum_{i=1}^{r} X_{(i)}.$$

   Show that $E(U) = \alpha^{-1}$ and $Var(U) = \alpha^{-2}/r$.

7. (Review MLE and asymptotic analysis) Let $X_1, \ldots, X_n$ be i.i.d. random variables with common density function

$$f_\theta(x) = \frac{\theta}{(1+x)^{\theta+1}}, \quad x > 0, \theta > 0.$$

   (a) Find the maximum likelihood estimator of θ, denoted as $\hat{\theta}_n$.

   (b) Find the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$.

(c) Find a function g such that, regardless the value of θ, $\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \to N(0,1)$ in distribution.

item (Law of large numbers and importance sampling) The following algorithm has been proposed to find $P(Z > 5) = 1 - \Phi(5)$ where Z is a standard normal random variable.

Step 1. Sample $x_1, x_2, \ldots, x_m$ from $N(5,1)$.

Step 2. Calculate

$$\hat{p}_{IS} = \frac{1}{m}\sum_{i=1}^{m} I(x_i \geq 5)\frac{\phi(x_i)}{\phi(x_i - 5)}$$

where $\phi(\cdot)$ is the density function of Z.

(a) Is $\hat{p}_{IS}$ an unbiased estimator of $1 - \Phi(5)$? Justify your answer.

(b) Determine the variance of $\hat{p}_{IS}$.

8. (Truncated random variables) Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$. For a fixed number x_0, define the function

$$g(x) = \begin{cases} f(x)/[1 - F(x_0)] & \text{if } x \geq x_0 \\ 0 & \text{if } x < x_0 \end{cases}$$

Prove that $g(x)$ is a pdf. (Assume that $F(x_0) < 1$.) Derive, via integration, the CDF (cumulative distribution function) corresponding to the pdf $g(x)$.

9. (Markovs inequality) Let X be a non-negative random variable and let g be an increasing non-negative function defined on $[0, \infty)$. Suppose that $E(g(X))$ is finite. Prove that, for any $\epsilon > 0$,

$$P(X \geq \epsilon) \leq E(g(X))/g(\epsilon).$$

10. (constrained optimization) Let X be a single observation from the $N(\theta, 1)$ distribution, $\theta > 0$. Notice that the unknown mean is assumed to be positive.

11

(4.)

(a) Obtain the maximum likelihood estimator (MLE) of θ, $\hat{\theta}$. (Be sure to consider the parameter space.)

(b) Show that $E(\hat{\theta}) \neq \theta$.

11. It has been determined that the total food consumption for a particular ant species in a region is a random variable. T which, given the number of colonies N, has $Normal(N\mu, N\sigma^2)$ distribution. Suppose $N \sim Poisson(\lambda)$. Suppose iid data $(T_1, N_1), (T_2, N_2), \ldots$ can be observed according to the above and let $\bar{T}$ and $\bar{N}$ be the corresponding means for samples of size n. Show that $\sqrt{n}(\bar{T} - \lambda\mu)$ and $\frac{\sqrt{n}}{\bar{N}}(\bar{T} - \lambda\mu)$ each converge in distribution, as $n \to \infty$, and describe the limits.

12. (Sample selection bias) Assume that ϵ and ν are random variables jointly normally distributed $N(0, 0, 1, 1, \rho)$, where $0 < \rho < 1$. Assume that a wage offer W are a random linear function of schooling S:

$$\begin{pmatrix}\epsilon \\ \nu\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho \\ \rho & 1\end{pmatrix}\right) \quad W = \alpha S + \epsilon.$$

Assume that wage offers are accepted (and therefore observed) only if the worker decides to enter the workplace. ν is a measure of *propensity to work*, and a worker enters the workplace only if $\nu > 0$. Derive an expression for the expected wage for the working population. Do not solve any complicated integrals.

13. Suppose U and V are independent with exponential distribution with parameter λ. (A random variable T is exponentially distributed with parameter λ if its density is given by $f(t) = \lambda \exp(-\lambda t)$ with support $T > 0$.) Define $X = U + V$ and $Y = UV$.

    (a) Derive the joint density of (X, Y).

    (b) Find the best linear predictor of Y given X.

    (c) Find the best predictor of Y given X.

14. Adapted from *A Simple Population Estimate Based on Simulation for Capture-Recapture and Capture-Resight Data* by Minta and Mangel, Ecology 1989.

    > In the fall of 1984, North American badgers (Taxidea taxus) were snowtracked in a 15 km^2 area on the National Elk Refuge, Jackson, Wyoming. The size and shape of the target area were dictated by topographic and plant community features that created a relatively isolated area of high badger density. Fifteen of the badgers were radiotagged and known to be occupying or overlapping the area. During the 2-month tracking period there was no death or emigration of radiotagged badgers, and radiotagged badgers outside the target area did not immigrate. One badger emigrated near the end of the sampling period. During daylight and under suitable weather conditions, the target was searched for badger snowtracks. A total of 24 tracks could be followed to a terminal hole, where the badger would be inactive in an underground burrow. All telemetry frequencies were then scanned to determine whether the badger was *marked* or *unmarked*. Radiotelemetry revealed that 11 of the tracks were generated by marked badgers.

    (a) Let N be the (unknown) total badger population size. Explain why the hypergeometric distribution can be used to model this experiment. Identify the values of the parameters M and K.

    (b) For the values of the parameters given above, what is your best guess (estimate) of N.

    (c) For the value of N from part (b), draw the hypergeometric distribution. How likely is the observed value of x?

    (d) For values of N near that of part (b), evaluate the probability of the observed value of x. What might you conclude about the population size?

作業3

高等統計推論 II

3

Advanced Statistical Inference II

Homework 2: Review of probability language, asymptotic analysis and some concepts in statistics.
Due Date: ~~March 20th~~, 2017

1. (root-finding algorithm) Let X_1, X_2, X_3 be independent observations from the Cauchy distribution about θ, $f(x,\theta) = \pi^{-1}(1+(x-\theta)^2)^{-1}$. Suppose $X_1 = 0$, $X_2 = 1$, $X_3 = a$. Show that for a sufficiently large the likelihood function has local maxima between 0 and 1 and θ and a. Deduce that depending on where bisection is started the sequence of iterates may converge to one or the other of the local maxima.

   不會

2. (convex function) Consider a convex function $f(x)$ where $x \in R$.

   Prove **(a)** If f is continuous, ~~then $f(x+y) \leq f(x) + f(y)$, for all $x, y \in R$~~.

   (b) If f is continuously differentiable, $f(y) \geq f(x) + f'(x)(y-x)$, for all $x, y \in R$.

   (c) If f is twice differentiable, $f''(x) \geq 0$, for all $x \in R$.

   Definition of convex set:
   A set $C \subset R^n$ is called convex set ,iffor all $x, y \in C$, $tx + (1-t)y \in C$ for $0 \leq t \leq 1$. That means for every pairs of points in the convex set C, every point falls on the straight line segment that joins the pair of points will also be in the set. A set that has such properties is super powerful, because you can get to any point in the set from a given point through a straight line without hitting the boundary. Hitting the boundary often means you might get stuck at a local minimum and never have the chance to find the global optimal solution.
   Definition of convex function:
   A function $f : R^p \to R$ is a convex function, if $dom(f)$ is a convex set, and $f(tx+(1-t)y) \leq tf(x) + (1-t)f(y)$ for $0 \leq tt \leq t1$. The inequalities say that the line segment between two points on the function will always lies above the function. Function that has such property is very powerful when you are trying to find the minimum value on the function. Because for any randomly chosen two points, x, y, you can always find a point z that lies between x and y such that $f(z) = \min(f(x), f(y))$. So you will surely be able find the optimal value.

3. (mle) An exponential distribution with parameter λ follows a distribution $p(x) = \lambda \exp(-\lambda x)$. Given some i.i.d. data $x_1, \cdots, x_n$ from $EXP(\lambda)$, derive the maximum likelihood estimate $\hat{\lambda}_{mle}$.

   (a) Is this estimator biased?

   (b) Is it consistent?

   (c) Determine its asymptotic distribution.

4. (Non-closed Form Estimation) Given some data $x_1, x_2, \cdots, x_3$ which is a random sample from $Gamma(\alpha, \beta)$, Use *gradient descent* to find the maximizer of the likelihood function and derive the steps to calculate the MLE estimators $(\hat{\alpha}_{mle}, \hat{\beta}_{mle})$. Please refer to the following webpage on *gradient descent*.

   `https://en.wikipedia.org/wiki/Gradient_descent`

5. Let X have density function $f(x|\sigma) = (2\sigma)^{-1}\exp(-|x|/\sigma)$. We have a random sample of size n from $f(\cdot)$. Derive the estimate of σ by the method of moments (denoted by $\hat{\sigma}_{mm}$) and maximum likelihood estimate (denoted by $\hat{\sigma}_{mle}$). Derive their asymptotic distrubutions.

6. (discrete random variable) Suppose we take one observation, X from a discrete random variable. X can take on five possible values on -2, -1, 0, 1, and 2 with probability $(1-\theta)/4$, $\theta/12$, $1/2$, $(3-\theta)/12$, $\theta/4$, respectively. Here, $0 \le \theta \le 1$.

(a) Find an unbiased estimator of θ.

(b) Obtain the maximum likelihood estimator (MLE) and show that it is not unique. Is any choice of MLE unbiased?

7. (order statistic) Let $X_1, \cdots, X_n$ be a random sample from the uniform distribution on the interval $(\theta-1, \theta+1)$, where $\theta \in R$ is unknown. Let $X_{(j)}$ be the jth order statistic.

(a) Show that $(X(1)+X(n))/2$ is strongly consistent for θ, i.e., that $\lim_{n\to\infty}(\cancel{X(1)+\cdots+X(n)})/2 = \theta$ almost ssurely.

(b) Show that $\bar{X}_n = (X_{(1)} + \cdots + X_{(n)})/n$ is L^2 consistent.

8. (exponential distribution) Let X be a random variable having probability density $f(x|\theta) = h(x)\exp[\eta(\theta)T(x) - A(\theta)]$, where η is an increasing and differentiable function of $\theta \in \Theta \subset R$.

(a) Show that $\log \ell(\hat\theta) - \log \ell(\theta_0)$ is increasing(or decreasing) in X, when $\hat\theta > \theta_0$ (or $\hat\theta < \theta_0$). Here $\ell(\theta) = f(x|\theta)$, $\hat\theta$ is an MLE of θ and $\theta_0 \in \Theta$.

(b) Show that $A(\theta) = \log \int (h(x)\exp(\theta T(x))dx$.
$\eta(\theta)$

9. (number of unknown parametrs is half of the sample size) Let X_{ij}, $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, k$ be independent with $X_{ij} \sim N(\mu_i, \sigma^2)$. Note that this is basically a balanced one-way ANOVA design where we assume k is fixed and $n \to \infty$. So the sample sizes of the groups are (probably) big, but the number of groups is bigger.

(a) Find the mle of μ_i and σ^2.

(b) Show that the MLE of σ^2, $\hat\sigma^2$ that is not consistent.

(c) Determine c such that $c\hat\sigma^2$ is a consistent estimate of σ^2.

10. Suppose $X_1, \cdots, X_n$ are p-vectors uniformly distributed in the ball $B_r = \{x : \|x\|_2 \le r\}$; $r > 0$ is an unknown parameter. Find the MLE of r and its asymptotic distribution.

11. (qualitative information on unknown parameters) Suppose $X_1, \cdots, X_{m+n}$ are independent, with $X_1, \cdots, X_m \sim N(\mu_1, \sigma^2)$, $X_{m+1}, \cdots, X_{m+n} \sim N(\mu_2, \sigma^2)$, where $\mu_1 \le \mu_2$ and σ^2 are unknown.

(a) Find the MLE of (μ_1, μ_2). $(\hat\mu_1, \hat\mu_2)$

(b) Derive its asymptotic distribution when $\mu_1 < \mu_2$, $\mu_1 = \mu_2$.
たぶん $(\hat\mu_1, \hat\mu_2)$ のMLEの漸近分布を求める?

12. (survival function) Suppose $X_1, \cdots, X_n$ are iid Exp(λ). Find the MLE of the expected residual life $E(X_1 - t|X_1 > t)$ and its asymptotic distribution.

$P(A \cap B) - P(A)P(B) =$

作業4

Advanced Statistical Inference II

Homework 4: Review of probability language, asymptotic analysis and some concepts in statistics.
Due Date: April ?, 2017

1. Suppose we take one observation, X, from a discrete distribution. It takes on values -2, -1, 0, 1, 2 with probabilites $(1-\theta)/4$, $\theta/12$, $1/2$, $(3-\theta)/12$, $\theta/4$, respectively. Her, $0 \leq \theta \leq 1$. Find an unbiased estimator of θ and obtain the maximum likelihood estimator (MLE) of θ and show that it is not unique. Is any choice of MLE unbiased?

2. Let $X_1, \cdots, X_n$ be a random sample of binary random variables with $P(X_1 = 1) = p$, where $p \in (0, 1)$ is unknown. Let $\hat{\theta}$ be the MLE of $\theta = p(1-p)$.

   (a) Show that $\hat{\theta}$ is asymptotically normal when $p \neq 1/2$.

   (b) When $p = 1/2$, derive a nondegenerated asymptotic distribution of $\hat{\theta}$ with an appropriate normalization.

3. Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be independent and identically distributed random 2-vectors taking values in the unit square $[0, 1] \times [0, 1]$, with joint CDF

$$F(x, y) = P(X_1 \leq x, Y_1 \leq y) = x \cdot y \cdot (x \wedge y)^{\theta}$$

   for $0 \leq x \leq 1$, $0 \leq y \leq 1$, where $\theta \geq 0$ is unknown.

   (a) Obtain the likelihood function for θ.

   (b) Obtain the MLE $\hat{\theta}_n((x_1, y_1), \cdots, (x_n, y_n))$ of θ.

   (c) Can you find the asymptotic distribution of $\hat{\theta}_n$? Hints: You can avoid the min by considering $x < y$ and $x \geq y$ separately; It is a little easier to work with the $\min U := (X \wedge Y)$ and $\max V := (X \wedge Y)$. Note that $F(x, y)$ is twice differentiable (once w.r.t x, once w.r.t. y) where $x \neq y$ but not on the diagonal where $x = y$ - in fact, $P(X = Y) > 0$, so you may want to consider observations with $X_i \neq Y_i$ and those with $X_i = Y_i$ separately. How can you find the p.d.f. for X when $X = Y$?

4. Let X be a random variable having probability density $f(x|\theta) = \exp\{\eta(\theta)Y(x) - B(\theta)\}h(x)$, where η is an increasing and differentiable function of $\theta \in \Theta \subset R$. Show that $\log \ell(\hat{\theta}) - \log \ell(\theta_0)$ is increasing (resp., decreasing) in Y, when $\hat{\theta} > \theta_0$ (resp., $\hat{\theta} < \theta_0$), where $\ell(\theta) = f(x|\theta)$, $\hat{\theta}$ is an MLE of θ, and $\theta_0 \in \Theta$.

5. (Review MLE and asymptotic analysis) Let $X_1, \ldots, X_n$ be i.i.d. random variables with common density function

$$f_\theta(x) = \frac{\theta}{(1+x)^{\theta+1}}, \quad x > 0, \theta > 0.$$

   (a) Find the maximum likelihood estimator of θ, denoted as $\hat{\theta}_n$.

   (b) Find the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$.

   (c) Find a function g such that, regardless the value of θ, $\sqrt{n}(g(\hat{\theta}_n) - g(\theta)) \to N(0, 1)$ in distribution.

6. Long (1990, 1997) gave data on the number of publications by 915 doctoral candidates in biochemistry. The numbers of published papers during last three years of Ph.D. are $0, 1, 2, 3, \cdots, 12$, 16 and 19. The corresponding number of students from 275, 246, 178, 84, 67, 27, 17, 12, 1, 2, 1, 1, 2, 1, 1. When the number of publications can be modelled as a Poisson random variable with mean λ, can you be certain that they come from Poisson based on the fact that $E(X) = Var(X) = \lambda$.

7. The double-exponential distribution has pdf $f(x) = (\lambda/2)\exp(-\lambda|x|)$, for fixed $\lambda > 0$.

   (a) Find the MGF $M_X(t)$ of a double exponential. For which t is it finite?

   (b) Let U and V are independent and identically distributed random variables with exponential distribution with mean 1,and 0find the MGF $M_Y(t)$ of $Y = U - V$. What is the distribution of Y?

   (c) Find the mean and variance of a double exponential. If $X_1, \cdots, X_n$ are i.i.d. double exponentials, find the MGF $M_n(t)$ of the standardized mean,

$$W_n = \frac{\bar{X} - E[X_1]}{\sqrt{Var(X)/n}}.$$

   (d) What is the limit of $M_n(t)$ as $n \to \infty$? What distribution has this function for its MGF?

8. Let $X_1, \cdots, X_n$ be IID normal random variables with mean μ and variance 1. Here $\mu \geq 1$.

   (a) Find the MLE of μ which is denoted by $\hat{\mu}_{mle}$.

   (b) When the true $\mu > 1$, determine the distribution of $\sqrt{n}(\hat{\mu}_{mle} - \mu)$ when n is large.

   (c) When the true $\mu = 1$, determine the distribution of $\sqrt{n}(\hat{\mu}_{mle} - \mu)$ when n is large.

9. Let $X_1, X_2 \sim Uniform(0, \theta)$ where $\theta > 0$.

   (a) Find the distribution of (X_1, X_2) given T where $T = \max\{X_1, X_2\}$.

   (b) Show that $X_1 + X_2$ is not sufficient.

10. Let $X_1, \cdots, X_n$ be a random sample from the density

$$f(x|\theta) = \theta x^{\theta - 1} I_{(0,1)}(x).$$

   The parameter space is $\Theta = (0, \infty)$.

   (a) Verify that $-\log X_1 = Y$ has an exponential distribution.

   (b) Find the Cramer Rao lower bound for unbiased estimators of $\tau(\theta) = 1/\theta$.

   (c) Show that $-\sum_{i=1}^{n} \log X_i/n$ is an UMVUE of $1/\theta$.

作業5 高等統計推論 II HW5

(New)

Copy11

Advanced Statistical Inference II

Homework 5: Hypothese Testing

Due Date: June 6th, 2017 6/6

1. Let $X_1, \ldots, X_n$ be iid with probability density $f_\theta(x)$. Denote $\hat{\theta}$ as the MLE. Suppose we are interested in estimating some function of θ.

   (a) If $\eta = g(\theta)$ and g is strictly increasing (or decreasing). Determine the MLE of η. (Justify your answer.)

   (b) Usually, the MLE of η is defined to be $g(\hat{\theta})$. The reason is the following:
   Let w_η be the set of θ-values for which $g(\theta) = \eta$ and let
   $$M(\eta) = \sup_{\theta \in w_n} L(\theta),$$
   where $L(\theta)$ is the likelihood function.
   Then it claims that if $\hat{\theta}$ maximizes $L(\theta)$, $\hat{\eta}$ maximizes $M(\eta)$.
   Justify the above claim. (If you are interested in this issue, read Zehna, Annals of Math. Stat., 1966.)

2. A random sample of five values, $x_1, x_2, \cdots, x_5$, is taken from a population whose density function is $f(x|\lambda) = x\lambda^2 \exp(-\lambda x)$ where $x > 0$.

   (a) Use the Neyman Pearson Lemma to construct a test of
   $$H_0 : \lambda = 2 \quad \text{against} \quad H_1 : \lambda = 1$$
   such that $\alpha = 0.05$ while it minimizes $3\alpha + \beta$. Here α and β are the Type I and Type II error probabilities respectively.

   (b) Find the approximate power of the test in (a) and the approximate value of a for the test in (b). (The approximations should be as accurate as the chisquared tables allow.)

   (c) If the alternative hypothesis had been $H_1 : \lambda < 1$ would the test be uniformly most powerful?

3. A new drug is being proposed for the treatment of migraine headaches. Unfortunately some users in early tests of the drug have reported mild nausea as a side effect. The FDA will reject the drug if it thinks that more than 15% (i.e., 0.15) of the population would suffer from this side effect. In an experiment to test this side effect, 400 people who suffer from migraine headaches receive the new drug and 80 of them report nausea as a side effect.

   (a) Define the parameter of interest, giving appropriate notation and writing a sentence saying what it is.

   (b) Carry out the five steps of a hypothesis test to determine if the FDA should reject the drug. (In Step 1, you need to specify the null and alternative hypotheses: Use notation, not words. In Step 2, compute the test statistic. (Show your work)

   (c) Find the p-value.

   (d) Decide whether the result is statistically significant (i.e. make a conclusion about the hypotheses); use $\alpha = 0.05$.

(e) Report the conclusion in context.

4. The following data give the number of aying bomb hits recorded in each of 576 small areas of $0.25km^2$ in the south of London during World War II. The following Table gives the number of flying bomb hits on London

Number of hits in an area	0	1	2	3	4	5	≥ 6
Frequency	229	211	93	35	7	1	0

1 ~ 229
230 ~ 440
441 ~ 533
534 ~ 568
569 ~ 57

Propaganda broadcasts claimed that the weapon could be aimed accurately. If, however, this was not the case, the hits should be randomly distributed over the area and should therefore be Otted by a Poisson distribution. Is this the case?

5. It is claimed that a new treatment is more effective than the standard treatment for prolonging the lives of terminal cancer patients. The standard treatment has been in use for a long time, and from records in medical journals the mean survival period has been 4.2 years with a standard deviation of 1.1 years. The new treatment is administered to 80 patients, and their average duration of survival is calculated to be 4.5 years. Is the claim supported by these results? In your answer, you should include H_0 versus H_a, p value, and your conclusion.

6. The manager at Costello Drug Store assumes the companys employees are honest. However, there have been many shortages from the cash register lately. There is only one employee who could have taken money from the register during these periods. Realizing that the shortages might have resulted from the employee inadvertently giving incorrect change to customers, the employer does not know whether to forget the situation or accuse the employee of theft. In words, what are the null and alternative hypotheses? Explain your choices.

   (a) What constitutes a Type I error in this problem?

   (b) What is a Type II error? Which do you think is more serious? Explain.

7. Let $X_1, X_2, \cdots$ be independent and identically distributed random variables with probability density function

$$f_\theta(x) = \theta \exp(-\theta x) \quad \text{for } x > 0.$$

The distribution is exponential with unknown, but fixed, scale $\theta > 0$. Define, for $\theta > 0$, the function

$$\phi_\theta(x) = \frac{1}{\theta} - x \quad \text{for } x > 0,$$

and consider the Z-estimator $\hat{\theta}_n$ defined as a zero of the function

$$\Phi_n(\theta) = P_n\phi_\theta = \frac{1}{n}\sum_{i=1}^{n} \phi_\theta(X_i).$$

   (a) Using properties of the function Φ_n, show that $\hat{\theta}_n$ is well defined, i.e. Φ_n has exactly one point where it takes the value zero.

   (b) Show that $\Phi_n(\theta)$ converges in probability, for fixed $\theta > 0$.

   (c) Prove that $\hat{\theta}_n$ converges to θ in probability.

   (d) The random variables $\sqrt{n}(\hat{\theta}_n - \theta)$ are asymptotically normally distributed. What parameters do you expect for this asymptotic normal distribution?

(e) Derive the maximum likelihood estimator for θ. How does it relate to $\hat{\theta}_n$? Based on this and (d), give the Fisher information I_θ.

8. Let $X_1, \ldots, X_n$ be i.i.d. from the Poisson distribution $P(\theta)$ with an unknown $\theta > 0$. Find the bias and mean squared error of $T_n = (1 - a/n)^{n\bar{X}}$ as an estimator of $\exp(-a\theta)$, where $a \neq 0$ is a known constant.

9. Let $X_1, \cdots, X_n$ be a sample from the normal distribution with mean θ and variance θ.

   (a) Find a variance stabilizing transformation for the sample mean and construct a confidence interval for based on this.

   (b) Find the limit distribution of the sequence $\sqrt{n}(\cos(\bar{X}_n) - \cos(\theta))$. For which values of $\theta$is this distribution degenerate?

10. Let $\Theta = \{-1, 0, 1\}$. Let f_θ be the pdf of a $N(\theta, 1)$ random variable. Consider the null hypothesis $\theta = 0$ and the alternate hypothesis $\theta \in \{1, -1\}$. Suppose we have one sample $X \sim N(\theta, 1)$ and we wish to test H_0 against H_a. Show that there is no uniformly most powerful test at level $\alpha \in (0, 1)$.

11. Let $\Theta = \{0, 1, 2\}$. Let f_θ be the pdf of a $N(\theta, 1)$ random variable. Consider the null hypothesis $\theta = 0$ and the alternate hypothesis $\theta \in \{1, 2\}$. Suppose we have one sample $X \sim N(\theta, 1)$ and we wish to test H_0 against H_a. Show that there is a uniformly most powerful test at level $\alpha \in (0, 1)$.

12. Let X be a random variable having probability density $f(x|\theta) = \exp\{\eta(\theta)Y(x) - \xi(\theta)\}h(x)$, where η is an increasing and differentiable function of $\theta \in \Theta \subset R$.

   (a) Show that $\log \ell(\hat{\theta}) - \log \ell(\theta_0)$ is increasing(or decreasing) in Y, when $\hat{\theta} > \theta_0$ (or $\hat{\theta} < \theta_0$), where $\ell(\theta) = f(x|\theta)$, $\hat{\theta}$ is an MLE of θ, and $\theta_0 \in \Theta$.

   復 **(b)** For testing $H_0 : \theta_1 \leq \theta \leq \theta_2$ versus $H_1 : \theta < \theta_1$ or $\theta > \theta_2$ or for testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$, show that there is a likelihood ratio test whose rejection region is equivalent to $Y(X) < c_1$ or $Y(x) > c_2$ for some constants c_1 and c_2. 復 … 利用して $\frac{dY}{d\theta}$ ≧0 を導く

13. 復 ✱ Let F and G be two known cumulative distribution functions(CDFs) on R and X be a single observation from the CDF $\theta F(x) + (1 - \theta)G(x)$, where $\theta \in [0, 1]$ is unknown.

   (a) Find a UMP test of size α for testing $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$, where $\theta_0 \in [0, 1]$ is known.

   (b) Show that the test $T(X) = \alpha$ is a UMP test of size α for testing $H_0 : \theta \leq \theta_1$ or $\theta \geq \theta_2$ versus $H_1 : \theta_1 < \theta < \theta_2$, where $\theta_j \in [0, 1]$ is known, $j = 1, 2$, and $\theta_1 < \theta_2$.

   (c) Derive the likelihood ratio $\lambda(X)$ for $H_0 : \theta \leq \theta_1$ or $\theta \geq \theta_2$.

・F. G.. differentiable
・θ (a.s) ,
という仮定を入れるべきか、

14. A family of probability density functions $f(x|\theta)$ on R, indexed by $\theta \in \Theta \subset R$, is said to have a *monotone likelihood ratio* (MLR) if, for each $\theta_0 \neq \theta_1$, the ratio $f(x|\theta_1)/f(x|\theta_0)$ is monotonic in x. Assume that $\partial^2 \log f(x|\theta)/\partial\theta\partial x$ exists.

(a) Show that a family of density functions $\{f(x|\theta) : \theta \in \Theta \subset R\}$ has MLR in x is equivalent to one of the following conditions:
(1) $\partial^2 \log f(x|\theta)/\partial\theta\partial x \geq 0$ for all x and θ;
(2) $f(x|\theta)\, [\partial^2 \cancel{\log} f(x|\theta)/\partial\theta\partial x] \geq (\partial f(x|\theta)/\partial\theta)(\partial f(x|\theta)/\partial x)$ for all x and θ.

(b) Let $Z \sim N(\sqrt{\theta}, 1)$, then Z^2 has the non-central chi-square distribution $\chi_1^2(\theta)$. Show that $\{\chi_1^2(\theta)\}$ has MLR in x.

15. A family of probability distributions on R, indexed by $\theta \in \Theta \subset R$, is called *stochastically increasing* if their CDFs $\{F(x|\theta), \theta \in \Theta\}$ satisfy the following statement

$$\text{When } \theta_1 < \theta_2, \text{ we have } F(x|\theta_1) \geq F(x|\theta_2) \quad \text{for all } x \in R.$$

(Intuitively this says that larger values of the parameter θ are associated with larger values of the random variable X).

(a) Show that if a family of pdfs $\{f(x|\theta) : \theta \in \Theta\}$ has an MLR, then the corresponding family of CDFs is stochastically increasing in θ.

(b) Show that the converse of part (a) is false.

R05246013

高等統計概論 II　作業①　森元 俊成

$\boxed{1}$ $Z \overset{\text{def}}{=} X+Y$

$$\Pr(X=x, Z=z) = \Pr(X=x, Z-X=z-x)$$

$$= \Pr(X=x, Y=z-x) = \Pr(X=x)\cdot\Pr(Y=z-x)$$

$$= e^{-\mu}\cdot\frac{\mu^x}{x!}\cdot e^{-\lambda}\cdot\frac{\lambda^{z-x}}{(z-x)!} \quad (x \geq 0,\ z-x \geq 0)$$

$$\Pr(X=x \mid Z=z) = \frac{\Pr(X=x, Z=z)}{\Pr(Z=z)}$$

$$\left(\begin{array}{l} \because\ X+Y=Z \sim Po(\lambda+\mu) \\ \Pr(Z=z) = e^{-(\lambda+\mu)}\cdot\frac{(\lambda+\mu)^z}{z!} \end{array}\right)$$

$$= \frac{e^{-\mu}\cdot\frac{\mu^x}{x!}\cdot e^{-\lambda}\cdot\frac{\lambda^{z-x}}{(z-x)!}}{e^{-(\lambda+\mu)}\cdot\frac{(\lambda+\mu)^z}{z!}}$$

$$= \frac{z!}{x!(z-x)!}\left(\frac{\mu}{\lambda+\mu}\right)^x\cdot\left(1-\frac{\mu}{\lambda+\mu}\right)^{z-x}$$

由此可知 $X \mid X+Y=z \sim Bin\left(z, \frac{\mu}{\lambda+\mu}\right)$

$\therefore X \mid Z=n \sim Bin\left(n, \frac{\mu}{\lambda+\mu}\right)$

No
Date

$\boxed{2}$ $X、Y \overset{iid}{\sim} e(\lambda)$

① $Pr(X^3 \leq t) = 1 - Pr(X^3 > t) = 1 - Pr(X > t^{\frac{1}{3}})$

$Pr(X > x) = \int_x^\infty \lambda \exp(-\lambda t) dt = [-\exp(-\lambda t)]_x^\infty$

$= \exp(-\lambda x)$

$\therefore Pr(X > t^{\frac{1}{3}}) = \exp(-\lambda t^{\frac{1}{3}})$

$\therefore Pr(X^3 \leq t) = 1 - \exp(-\lambda t^{\frac{1}{3}})$

$\frac{d}{dt} Pr(X^3 \leq t) = \frac{\lambda t^{-\frac{2}{3}}}{3} \exp(-\lambda t^{\frac{1}{3}})$

$\therefore f_{X^3}(x) = \frac{\lambda x^{-\frac{2}{3}}}{3} \exp(-\lambda x^{\frac{1}{3}})$

② $|X-Y|$ … sample range

將情況一般化：

$X_1, X_2 \cdots X_n \overset{iid}{\sim} e(\lambda) \quad (n=2)$

$R = X_{(n)} - X_{(1)}$

$S = \frac{X_{(n)} + X_{(1)}}{2}$

③

No.
Date

$$\Pr(X_{(1)} \leqq x,\ X_{(n)} \leqq y)$$

$$= \Pr(X_{(n)} \leqq y) - \Pr(x < X_{(1)}, X_{(n)} \leqq y)$$

$$\Pr(X_1 \sim X_n \leqq y) \qquad \Pr(x < X_1 \sim X_n \leqq y)$$

$$\left(\int_0^y \lambda \exp(-\lambda t)\,dt\right)^n \qquad \left(\exp(-\lambda x) - \exp(-\lambda y)\right)^n$$

$$= \left(1 - \exp(-\lambda y)\right)^n$$

$$\therefore \frac{\partial^2}{\partial x \partial y} \Pr(X_{(1)} \leqq x,\ X_{(n)} \leqq y)$$

$$= n(n-1)\left(\exp(-\lambda x) - \exp(-\lambda y)\right)^{n-2} \cdot \lambda\exp(-\lambda x) \cdot \lambda\exp(-\lambda y)$$

$$= f_{X_{(1)}, X_{(n)}}(X_{(1)} = x,\ X_{(n)} = y)$$

$$\begin{cases} r = y - x \\ s = \dfrac{y+x}{2} \end{cases}$$ 變数轉換 $\begin{pmatrix} X_{(1)} \\ X_{(n)} \end{pmatrix} \to \begin{pmatrix} R \\ S \end{pmatrix}$

$$\Leftrightarrow \begin{cases} x = s - \dfrac{r}{2} \\ y = s + \dfrac{r}{2} \end{cases}$$

$$dx\,dy = dr\,ds$$

Chryu culture

No.
Date

$$\begin{array}{ccccc} & X_{(1)} & & X_{(2)} & \\ & \| & & \| & \\ 0 \leqq & s-\frac{r}{2} & \leqq & s+\frac{r}{2} & < \infty \end{array}$$

$\Leftrightarrow \frac{r}{2} \leqq s, \ r \geqq 0.$

$$f_{R,S}(r,s) = n(n-1)\left\{\exp\left(-\lambda\left(s-\frac{r}{2}\right)\right) - \exp\left(-\lambda\left(s+\frac{r}{2}\right)\right)\right\}^{n-2}$$

$$\cdot \lambda^2 \exp(-\lambda(2s))$$

$$= n(n-1)\ \exp(-\lambda s)^{(n-2)} \cdot \left\{\exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right)\right\}^{(n-2)}$$

$$\lambda^2 \exp(-2\lambda s)$$

$$= n(n-1)\lambda^2 \exp(-\lambda n s)\left\{\exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right)\right\}^{(n-2)}$$

$f_R(r) =$

$$\int_{s \geqq \frac{r}{2}} f_{R,S}(r,s)\,ds = \int_{s \geqq \frac{r}{2}} \lambda n \exp(-\lambda n s)\,ds$$

$$\cdot\ \lambda(n-1)\left\{\exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right)\right\}^{(n-2)}$$

$$= \Big[\exp(-\lambda n s)\Big]_{\frac{r}{2}}^{\infty} \cdot \lambda(n-1)\left\{\exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right)\right\}^{n-2}$$

$$= \exp\left(\frac{-\lambda n r}{2}\right) \cdot \lambda(n-1)\left\{\exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right)\right\}^{n-2}$$

$$= f_R(r)$$

$n=2$.. $f_R(r) = \lambda \exp(-\lambda r)$ $\quad \therefore |X-Y| \sim \exp(\lambda)$

$(r \geqq 0)$

⑤

No.

Date

③ $\min\{X, Y^3\}$

$$\Pr(\min\{X, Y^3\} \leq t) = 1 - \Pr(\min\{X, Y^3\} > t)$$

$$= 1 - \Pr(X > t)\Pr(Y^3 > t)$$

$$= 1 - \Pr(X > t)\Pr(Y > t^{\frac{1}{3}})$$

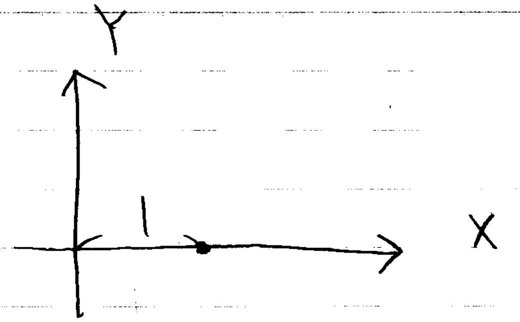
$$= 1 - \exp(-\lambda t)\exp(-\lambda t^{\frac{1}{3}})$$

$$= 1 - \exp(-\lambda(t + t^{\frac{1}{3}}))$$

$$\frac{d}{dt}(1 - \exp(-\lambda(t + t^{\frac{1}{3}}))) = \lambda(1 + \frac{1}{3}t^{-\frac{2}{3}})\exp(-\lambda(t + t^{\frac{1}{3}}))$$

$$\therefore f_{\min\{X, Y^3\}}(t) = \lambda(1 + \frac{1}{3}t^{-\frac{2}{3}})\exp(-\lambda(t + t^{\frac{1}{3}})) \quad (t \geq 0)$$

3



$$\theta \sim U(-\pi, \pi)$$

$$l = \| (\cos\theta, \sin\theta) - (1,0) \|$$

$$= \sqrt{(1-\cos\theta)^2 + \sin^2\theta} = \sqrt{2-2\cos\theta}$$

$$= \sqrt{2-2\cos\left(\frac{\theta}{2}+\frac{\theta}{2}\right)} = \left| 2\sin\frac{\theta}{2} \right|$$

∴ θ 之機率密度函數 $f(\theta) = \frac{1}{2\pi} I_{(-\pi,\pi)}(\theta)$

$$1 = \int_{-\pi}^{\pi} \frac{1}{2\pi} d\theta = \int_{0}^{\pi} \frac{1}{2\pi} d\theta + \int_{-\pi}^{0} \frac{1}{2\pi} d\theta$$

① $\theta: 0 \to \pi$ ⋯ $l = 2\sin\frac{\theta}{2}$ $\theta = 2\arcsin\left(\frac{l}{2}\right)$

$l: 0 \to 2$

$$\frac{d\theta}{dl} = \frac{1}{\sqrt{1-\left(\frac{l}{2}\right)^2}}$$

② $\theta: -\pi \to 0$ $l = -2\sin\frac{\theta}{2}$ $\theta = 2\arcsin\left(\frac{-l}{2}\right)$

$l: 2 \to 0$

$$\frac{d\theta}{dl} = \frac{(-1)}{\sqrt{1-\left(\frac{-l}{2}\right)^2}}$$

⑰

No

Date

$$\int_0^{\pi} \frac{1}{2\pi}\, d\theta + \int_{-\pi}^{0} \frac{1}{2\pi}\, d\theta$$

$$= \int_0^2 \frac{1}{2\pi} \cdot \frac{dl}{\sqrt{1-\frac{l^2}{4}}} + \int_2^0 \frac{1}{2\pi} \cdot \frac{-dl}{\sqrt{1-\frac{l^2}{4}}}$$

$$= \int_0^2 \frac{1}{\pi} \cdot \frac{dl}{\sqrt{1-\frac{l^2}{4}}}$$

$$\therefore f_L(l) = \frac{2}{\pi} \cdot \frac{1}{\sqrt{4-l^2}} \quad (0 < l < 2)$$

No.

Date

$\boxed{4}$ $Pr(Y \leq y) = Pr(F^{-1}(X) \leq y)$

(F: 遞增函數) $= Pr(X \leq F(y)) = F(y)$

由此可知，Y之 cdf 亦為 F.

$\therefore Y \sim F$

$\boxed{5}$ 首先考慮 $W_1, W_2, W_3, W_4 \overset{iid}{\sim} exp(1)$

$$\begin{cases} Y_1 \overset{def}{=} \dfrac{W_1}{W_1+W_2+W_3+W_4} \\ Y_2 \overset{def}{=} \dfrac{W_1+W_2}{W_1+W_2+W_3+W_4} \\ Y_3 \overset{def}{=} \dfrac{W_1+W_2+W_3}{W_1+W_2+W_3+W_4} \\ Y_4 \overset{def}{=} W_1+W_2+W_3+W_4 \end{cases}$$

$$\Rightarrow \begin{cases} W_1 = Y_1 Y_4 \\ W_2 = Y_2 Y_4 - Y_1 Y_4 \\ W_3 = Y_3 Y_4 - Y_2 Y_4 \\ W_4 = Y_4 - Y_3 Y_4 \end{cases}$$

9

No.
Date

$$\frac{\partial}{\partial y_1}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} y_4 \\ -y_4 \\ 0 \\ 0 \end{pmatrix} \quad \frac{\partial}{\partial y_2}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} 0 \\ y_4 \\ -y_4 \\ 0 \end{pmatrix}$$

$$\frac{\partial}{\partial y_3}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ y_4 \\ -y_4 \end{pmatrix} \quad \frac{\partial}{\partial y_4}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2-y_1 \\ y_3-y_2 \\ 1-y_3 \end{pmatrix}$$

$$\therefore J = \begin{pmatrix} y_4 & 0 & 0 & y_1 \\ -y_4 & y_4 & 0 & y_2-y_1 \\ 0 & -y_4 & y_4 & y_3-y_2 \\ 0 & 0 & -y_4 & 1-y_3 \end{pmatrix}$$

$$\det J = y_4^3 \cdot \det\begin{pmatrix} 1 & 0 & 0 & y_1 \\ -1 & 1 & 0 & y_2-y_1 \\ 0 & -1 & 1 & y_3-y_2 \\ 0 & 0 & -1 & 1-y_3 \end{pmatrix}$$

$$= y_4^3 \cdot \det\begin{pmatrix} 1 & 0 & 0 & y_1 \\ 0 & 1 & 0 & y_2 \\ 0 & 0 & 1 & y_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} = y_4^3$$

$$\therefore \quad dw_1\,dw_2\,dw_3\,dw_4 = y_4^3\,dy_1\,dy_2\,dy_3\,dy_4$$

$$\text{全確率} = 1 = \int_{\substack{w_1>0 \\ \vdots \\ w_4>0}} \exp(-(w_1+w_2+w_3+w_4))\,dw_1 \cdots dw_4$$

$$= \int_{\substack{0<y_1<y_2<y_3<1 \\ y_4>0}} y_4^3 \exp(-y_4)\,dy_4\,dy_1\,dy_2\,dy_3$$

Chryv culture

No
Date

$$f_{Y_1,Y_2,Y_3}(y_1,y_2,y_3) = \int_{y_4>0} y_4^3 \exp(-y_4)\,dy_4$$

$$= \Gamma(4)\, I(0<y_1<y_2<y_3<1)$$

$$= 3!\, I(0<y_1<y_2<y_3<1)$$

由此可知, $X_1, X_2, X_3 \overset{iid}{\sim} U(0,1)$

$$\Rightarrow \begin{pmatrix} X_{(1)} \\ X_{(2)} \\ X_{(3)} \end{pmatrix} \text{之分佈等於} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$$

(順序統計量)

根據題意, $U_1 = \dfrac{X_{(1)}}{X_{(2)}} = \dfrac{Y_1}{Y_2} = \dfrac{W_1}{W_1+W_2}$

$$U_2 = \frac{X_{(2)}}{X_{(3)}} = \frac{Y_2}{Y_3} = \frac{W_1+W_2}{W_1+W_2+W_3}$$

$(W_1, W_2, W_3 \overset{iid}{\sim} e(1))$

$$U_3 \overset{def}{=} W_1+W_2+W_3$$

$$\Rightarrow \begin{cases} W_1 = U_1U_2U_3 \\ W_2 = U_2U_3 - U_1U_2U_3 \\ W_3 = U_3 - U_2U_3. \end{cases}$$

⑪

No.
Date

$$\frac{\partial}{\partial u_1}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} u_2u_3 \\ -u_2u_3 \\ 0 \end{pmatrix} \quad \frac{\partial}{\partial u_2}\begin{pmatrix} u_1u_3 \\ u_3 - u_1u_3 \\ -u_3 \end{pmatrix}$$

$$\frac{\partial}{\partial u_3}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} u_1u_2 \\ u_2 - u_1u_2 \\ 1-u_2 \end{pmatrix}$$

$$\therefore J = \begin{pmatrix} u_2u_3 & u_1u_3 & u_1u_2 \\ -u_2u_3 & u_3-u_1u_3 & u_2-u_1u_2 \\ 0 & -u_3 & 1-u_2 \end{pmatrix}$$

$$\det J = u_2u_3^2 \cdot \det\begin{pmatrix} 1 & u_1 & u_1u_2 \\ -1 & 1-u_1 & u_2-u_1u_2 \\ 0 & -1 & 1-u_2 \end{pmatrix}$$

$$= u_2u_3^2 \det\begin{pmatrix} 1 & u_1 & u_1u_2 \\ 0 & 1 & u_2 \\ 0 & 0 & 1 \end{pmatrix} = u_2u_3^2$$

$$w_1 \geqq 0,\ w_2 \geqq 0,\ w_3 \geqq 0$$

$$\therefore\ 0 \leqq u_1 \leqq 1,\ 0 \leqq u_2 \leqq 1,\ u_3 \geqq 0$$

全確率 $= 1 = \displaystyle\int_{\substack{w_1>0 \\ w_2>0 \\ w_3>0}} \exp(-(w_1+w_2+w_3))\, dw_1 dw_2 dw_3$

$$= \int_{\substack{0\leqq u_1 \leqq 1 \\ 0 \leqq u_2 \leqq 1 \\ u_3 \geqq 0}} \exp(-u_3) \cdot u_2u_3^2\, du_1 du_2 du_3$$

Chryo culture

No.
Date

$$f_{U_1,U_2}(u_1,u_2) = \int_{u_3 \geq 0} u_2 \cdot u_3^2 \exp(-u_3)\,du_3$$

$$= \Gamma(3) \cdot u_2 = 2u_2$$

$$U_1 = \frac{W_1}{W_1+W_2} \qquad U_2 = \frac{W_1+W_2}{W_1+W_2+W_3}$$

利用以下事實… $\begin{pmatrix} X \sim \Gamma(\alpha,\lambda) \\ Y \sim \Gamma(\beta,\lambda) \end{pmatrix}$，X, Y 獨立

$$\Rightarrow \frac{X}{X+Y} \sim Be(\alpha,\beta)$$

根據這個事實，$U_1 = \frac{W_1}{W_1+W_2} \sim Be(1,1) = \mathrm{Uni}(0,1)$

$$U_2 = \frac{W_1+W_2}{W_1+W_2+W_3} \sim Be(2,1)$$

$$\therefore f_{U_1}(u_1) = 1 \quad (0 \leq u_1 \leq 1)$$

$$f_{U_2}(u_2) = \frac{1}{Be(2,1)} u_2^{2-1}(1-u_1)^{1-1} = 2u_2 \quad (0 \leq u_2 \leq 1)$$

由此可知，$f_{U_1U_2}(u_1,u_2) = f_{U_1}(u_1)\, f_{U_2}(u_2)$

$\therefore$ U_1, U_2 為獨立。

No
Date

[6] $X_1, X_2 \cdots X_n \sim \exp(\lambda)$

• 先求 $X_{(1)} \cdots X_{(n)}$ 之聯合機率密度函數

$$f_{X_1 \cdots X_n}(x_1 \cdots x_n) = \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_n)$$

n個順序統計量的聯合機率密度函數

$$f_{X_{(1)} \cdots X_{(n)}}(x_1 \cdots x_n) = n!\, f_{X_1 \cdots X_n}(x_1 \cdots x_n)$$
$$(0 \leq x_1 \leq \cdots \leq x_n)$$

• Step1 … $X_{(1)} \cdots X_{(n-1)}$ 之聯合分佈

$$n! \int_{x_n > x_{n-1}} \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_n)\, dx_n$$

$$= n!\, \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_{n-1}) \cdot \exp(-\lambda x_{n-1})$$

$$= n!\, \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_{n-2})\, \lambda \exp(-2\lambda x_{n-1})$$

• Step2 … $X_{(1)} \cdots X_{(n-2)}$ 之聯合分佈

$$n! \int_{x_{n-1} > x_{n-2}} \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_{n-2}) \cdot \lambda \exp(-2\lambda x_{n-1})\, dx_{n-1}$$

$$= \frac{n!}{2!} \cdot \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_{n-3})\ \lambda \exp(-3\lambda x_{n-2})$$

⋮

No

Date

- Step j… $X_{(1)} \dots X_{(n-j)}$ 之聯合分佈

$$= \frac{n!}{j!} \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_{n-j-1}) \lambda \exp(-(j+1)\lambda x_{n-j})$$

$j = n-r$ 得到 $X_{(1)} \dots X_{(r)}$ 之聯合分佈

$$\because \frac{n!}{(n-r)!} \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_{r-1}) \lambda \exp(-(n-r+1)\lambda x_r)$$

$$(x_1 \leq x_2 \leq \cdots x_r)$$

接下來考慮以下變數轉換

$$\begin{cases} Y_1 = nX_{(1)} \\ Y_2 = (n-1)(X_{(2)} - X_{(1)}) \\ \vdots \\ Y_r = (n-r+1)(X_{(r)} - X_{(r-1)}) \end{cases}$$

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_r \end{pmatrix} = \begin{pmatrix} n & & & \\ -(n-1) & (n-1) & & \\ & -(n-2) & (n-2) & \\ & & \ddots & \\ & & & (n-r+1) \end{pmatrix} \begin{pmatrix} X_{(1)} \\ \vdots \\ X_{(r)} \end{pmatrix}$$

$$dy_1 \cdots dy_r = \frac{n!}{(n-r)!} dx_1 \cdots dx_n$$

⑮

No.
Date :

$X_{(1)} \leq X_{(2)} \cdots \leq X_{(r)}$

$\therefore Y_1 \geq 0,\ Y_2 \geq 0 \cdots\ Y_r \geq 0$

$Y_1, Y_2 \cdots Y_r$ 之聯合機率密度函數

全機率 $= 1 = \int_{0 \leq x_1 \cdots \leq x_r} \frac{n!}{(n-r)!} \lambda \exp(-\lambda x_1) \cdots \lambda \exp(-\lambda x_{r-1})$

$\lambda \cdot \exp(-(n-r+1)\lambda x_r)\, dx_1 \cdots dx_r$

$= \int_{y_i > 0} \lambda^r \exp(-\lambda(y_1 + \cdots + y_r))\, dy_1 \cdots dy_r$

$f_{Y_1 \cdots Y_r}(y_1 \cdots y_r) = \lambda^r \exp(-\lambda(y_1 + \cdots y_r)) \quad (y_1 \sim y_r > 0)$

$f_{Y_i} = \lambda \exp(-\lambda y_i)$

$f_{Y_1 \cdots Y_r}(y_1 \cdots y_r) = f_{Y_1}(y_1) \cdots f_{Y_r}(y_r) \quad (Y_1 \sim Y_r$ 獨立$)$

由此可知, $Y_1, Y_2 \cdots Y_r \overset{iid}{\sim} \exp(\lambda)$

$rU = (n-r) X_{(r)} + \sum_{i=1}^{r} X_{(i)}$

$= Y_1 + Y_2 + \cdots + Y_r \sim \Gamma(r, \lambda)$ (Gamma分布)

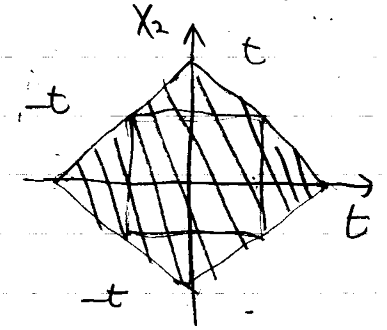
$\therefore E[rU] = \frac{r}{\lambda} \quad V[rU] = \frac{r}{\lambda^2} \Rightarrow$

$\therefore E[U] = \frac{1}{\lambda} \quad V[U] = \frac{1}{r\lambda^2}$

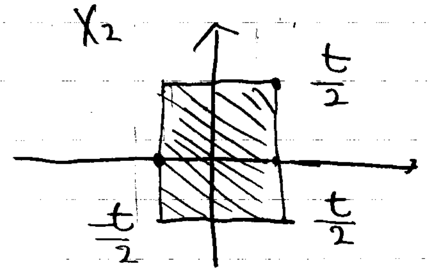
$\boxed{17}$ $A=\{(X_1,X_2)\mid |X_1-X_2|>t\}$

$B=\{(X_1,X_2)\mid |X_1|>\frac{t}{2}\}\cup\{(X_1,X_2)\mid |X_2|>\frac{t}{2}\}$

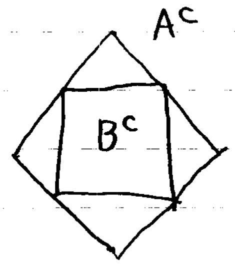
A^c ……



B^c …



由此可知, $A^c \supset B^c$



$\therefore A \subseteq B$

$\therefore Pr(A) \leq Pr(B)$

$= Pr(\{(X_1,X_2)\mid |X_1|>\frac{t}{2}\}\cup\{(X_1,X_2)\mid |X_2|>\frac{t}{2}\})$

$= 2Pr(|X_1|>\frac{t}{2})$

$(\because X_1, X_2 \cdots \text{i.i.d})$

證明完成

⑰

No.
Date

8

(a) $f_\theta(x_1 \dots x_n) = \dfrac{\theta}{(1+x_1)^{\theta+1}} \cdots \dfrac{\theta}{(1+x_n)^{\theta+1}}$

$\log f_\theta(x_1 \dots x_n) = n\log\theta - (\theta+1)\sum_{j=1}^{n} \log(1+x_j)$

$\dfrac{\partial}{\partial\theta}\log f_\theta(x_1 \dots x_n) = \dfrac{n}{\theta} - \sum_{j=1}^{n}\log(1+x_j)$

$\dfrac{\partial^2}{\partial\theta^2}\log f_\theta(x_1 \dots x_n) = \dfrac{-n}{\theta^2} < 0$

$\dfrac{\partial}{\partial\theta}\log f_\theta(x_1 \dots x_n) = 0$

$\Rightarrow \theta = \dfrac{n}{\sum_{j=1}^{n}\log(1+x_j)}$

$\therefore \hat{\theta}_{MLE} = \dfrac{n}{\sum_{j=1}^{n}\log(1+x_j)}$

(b) 流程… 求 $\log(1+X_j)$ 之分佈

↓

中央極限定理

↓

δ - method

No.
Date

$t = \log(1+x) \qquad (1+x) = e^{t} \qquad \begin{pmatrix} x: 0 \to \infty \\ t: 0 \to \infty \end{pmatrix}$

$\frac{dt}{dx} = \frac{1}{1+x}$

$$全確率 = 1 = \int_{x>0} \frac{\theta}{(1+x)^{\theta+1}} dx$$

$$= \int_{t>0} \theta \cdot e^{-\theta t} \cdot dt$$

$\therefore T_j = \log(1+X_j) \sim \exp(\theta) \qquad \left(E[T] = \frac{1}{\theta} \quad V[T] = \frac{1}{\theta^2} \right)$

$E[T] < \infty, \quad E[T^2] < \infty$

根據中央極限定理，

$$\sqrt{n}\left(\frac{T_1 + \cdots T_n}{n} - \frac{1}{\theta}\right) \xrightarrow{d} N\left(0, \frac{1}{\theta^2}\right)$$

接下來利用 δ-method

$g(x) = \frac{1}{x}$

$$\sqrt{n}\left(g\left(\frac{T_1 + \cdots T_n}{n}\right) - g\left(\frac{1}{\theta}\right)\right) \xrightarrow{d} N\left(0, \frac{1}{\theta^2}\left(g'\left(\frac{1}{\theta}\right)\right)^2\right)$$

$g'(x) = \frac{-1}{x^2} \qquad g'\left(\frac{1}{\theta}\right) = -\theta^2 \qquad \left(g'\left(\frac{1}{\theta}\right)\right)^2 = \theta^4$

19

No.

Date

$\therefore \sqrt{n}\left(g\left(\frac{X_1+\cdots+X_n}{n}\right)-g\left(\frac{1}{\theta}\right)\right) \xrightarrow{d} N(0,\theta^2)$

$\sqrt{n}(\hat{\theta}_{MLE}-\theta) \xrightarrow{d} N(0,\theta^2)$

(c) 同樣利用 δ-method

$\sqrt{n}\left(g\left(\frac{X_1+\cdots+X_n}{n}\right)-g\left(\frac{1}{\theta}\right)\right) \xrightarrow{d} N\left(0, \frac{1}{\theta^2}\cdot\left(g'\left(\frac{1}{\theta}\right)\right)^2\right)$

$\therefore g'\left(\frac{1}{\theta}\right)=\pm\theta$

$\therefore g'(\theta)=\pm\frac{1}{\theta}$

$g(\theta)=\pm\log\theta+C \quad (C=0)$

在此採用 $g(\theta)=-\log\theta$

$\sqrt{n}\left(-\log\left(\frac{X_1+\cdots+X_n}{n}\right)+\log\left(\frac{1}{\theta}\right)\right) \xrightarrow{d} N(0,1)$

$\Leftrightarrow \sqrt{n}(\log\hat{\theta}_{MLE}-\log\theta) \xrightarrow{d} N(0,1)$

(同樣道理 $\log\theta$ 亦可)

$g(x)=\pm\log x+C \quad (C:$ 常數$)$

No
Date

9

(a) $E[\hat{p}] = E\left[\frac{1}{m}\sum_{i=1}^{m} I(x_i \geq 5)\cdot\frac{\phi(x_i)}{\phi(x_i-5)}\right]$

$= \frac{1}{m}\sum_{i=1}^{m} E\left[I(X_i \geq 5)\cdot\frac{\phi(X_i)}{\phi(X_i-5)}\right]$

$= \frac{1}{m}\sum_{i=1}^{m}\int_{x_i\in\mathbb{R}} I(X_i \geq 5)\cdot\frac{\phi(x_i)}{\phi(x_i-5)}\cdot\phi(x_i-5)\,dx_i$

$= \frac{1}{m}\sum_{i=1}^{m}\int_{x_i\geq 5}\phi(x_i)\,dx_i$

$= \frac{1}{m}\sum_{i=1}^{m}(1-\Phi(5)) = 1-\Phi(5)$

$\therefore \hat{p}$ 為 $1-\Phi(5)$ 之不偏估計量

(b) $V[\hat{p}] = V\left[\frac{1}{m}\sum_{i=1}^{m} I(X_i \geq 5)\frac{\phi(X_i)}{\phi(X_i-5)}\right]$

$= \frac{1}{m}V\left[I(X_i\geq 5)\cdot\frac{\phi(X_i)}{\phi(X_i-5)}\right]$

- $E\left[\left(I(X_i\geq 5)\cdot\frac{\phi(X_i)}{\phi(X_i-5)}\right)\right] = 1-\Phi(5)$
- $E\left[\left(I(X_i\geq 5)\cdot\frac{\phi(X_i)}{\phi(X_i-5)}\right)^2\right] = \int_{x_i\geq 5}\frac{\phi(X_i)^2}{\phi(X_i-5)^2}\cdot\phi(X_i-5)\,dx_i$

$= \int_{x\geq 5}\frac{\phi(x)^2}{\phi(x-5)}\,dx$

(21)

No.

Date

$$\int_{x\geq 5} \frac{\frac{1}{2\pi}\exp(-x^2)}{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x-5)^2\right)}\,dx$$

$$=\int_{x\geq 5} \frac{1}{\sqrt{2\pi}}\exp\left(-x^2+\frac{1}{2}(x-5)^2\right)dx$$

$$=\int_{x\geq 5} \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-x^2-10x+25}{2}\right)dx$$

$$=\int_{x\geq 5} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x+5)^2+\frac{50}{2}\right)dx$$

$$y=x+5$$

$$=\int_{y\geq 10} \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-y^2}{2}\right)\cdot\exp(25)\,dy$$

$$=\exp(25)\cdot(1-\Phi(10))$$

$$\therefore V\left[I(X_i\geq 5)\cdot\frac{\phi(X_i)}{\phi(X_i-5)}\right]$$

$$=\exp(25)(1-\Phi(10))-(1-\Phi(5))^2$$

$$\therefore V[\hat{p}]=\frac{1}{m}\left\{\exp(25)(1-\Phi(10))-(1-\Phi(5))^2\right\}$$

No

Date

$\boxed{10}$ $Z(\lambda,\nu)$ 應為 $\sum_{i=0}^{\infty}\frac{\lambda^i}{(i!)^\nu}$ $(i=1 \to i=0)$

(a) $\Pr(Y=y \mid \lambda,\nu)=\frac{\lambda^y}{(y!)^\nu}\cdot\frac{1}{Z(\lambda,\nu)}$

$= \exp\left(\log \frac{\lambda^y}{(y!)^\nu}\cdot\frac{1}{Z(\lambda,\nu)}\right)$

$= \exp\left(y\log\lambda - \nu\log y! - \log Z(\lambda,\nu)\right)$

$= \exp\left(\sum_{j=1}^{2} C_j(\lambda,\nu)T_j(y) + d(\lambda,\nu)\right)$

$C_1(\lambda,\nu)=\log\lambda$, $T_1(y)=y$, $d(\lambda,\nu)=-\log Z(\lambda,\nu)$

$C_2(\lambda,\nu)=-\nu$ $T_2(y)=y!$

∴ 此分佈為指數族.

(b) $\nu=1 \Rightarrow Z(\lambda,\nu)=\sum_{i=0}^{\infty}\frac{\lambda^i}{(i!)}=1+\lambda+\frac{\lambda^2}{2!}+\cdots=e^\lambda$

$\Pr(Y=y)= e^{-\lambda}\cdot\frac{\lambda^y}{y!}$

∴ $Y\sim Po(\lambda)$

∴ $\nu=1$ 時 $Y\sim Po(\lambda)$

(C) Poisson分佈的期望值與變異數一致
$(E[X]=V[X]=\lambda)$
但實際上觀測到的樣本平均與
樣本變異數可能有很大的差異.
在這種情況下, Poisson分布不適合用來解釋
該筆資料的分布。
Conway-Maxwell-Poisson分布, 根據ν之值,
除了Poisson分布之外, 也可以表達幾何分布
Bernulli分布等的分布。
利用這個模型可以給更自然的解釋。

$\boxed{11}$ 在此，將題意解釋為利用 χ^2-test 檢定資料是否來自 Poisson 分布。（H_0: 資料來自 Poisson 分佈）

$$\hat{\lambda} = \frac{1}{n}\sum_{j=1}^{n} x_j = \underset{\sim\sim\sim}{1.692\cdots} \quad (n=915)$$

$$T = \sum_{j=0}^{k} \frac{(E_i - O_i)^2}{E_i} \quad (k=13,\ \text{將}(\geq 13\ \text{整理在一起})$$

j	0	1	2		≥ 13
E_i	168.3491	284.8494	⋯	⋮	$2.881 \cdot 10^{-5}$
O_i	275	246	⋯		2

$$T \sim \chi^2_{\underset{\downarrow}{14-1-1}} = \chi^2_{12}$$

（估計1個參數，所以減一個自由度）

$$T = 1604963 \gg \chi^2_{12}(0.95)$$

棄却 H_0.

25

12 $X \sim \text{Laplace}(\lambda, \mu=0)$

(a)
$$f(x) = \frac{\lambda}{2}\exp(-\lambda|x|)$$

$$M_X(t) = \int_R e^{tx} \cdot \frac{\lambda}{2}\exp(-\lambda|x|)\,dx$$

$$= \int_{x\geq 0} e^{tx} \cdot \frac{\lambda}{2}\exp(-\lambda x)\,dx$$

$$+ \int_{x<0} e^{tx} \cdot \frac{\lambda}{2}\exp(\lambda x)\,dx \quad \cdots (x \to -x)$$

$$= \frac{1}{2}\int_{x\geq 0} \lambda e^{tx} \cdot e^{-\lambda x}\,dx \qquad (\lambda - t > 0)$$

$$+ \frac{1}{2}\int_{x\geq 0} \lambda e^{-tx} \cdot e^{-\lambda x}\,dx \qquad (\lambda + t > 0)$$

$Y \sim \exp(\lambda)$

$$M_Y(t) = \int_{y\geq 0} \lambda e^{ty} \cdot e^{-\lambda y}\,dy \qquad (\lambda - t > 0)$$

$$= \int_{y\geq 0} \lambda e^{-(\lambda - t)y}\,dy$$

$$= \left[\frac{-\lambda}{\lambda - t} e^{-(\lambda - t)y}\right]_{y=0}^{y\to\infty}$$

$$= \frac{\lambda}{\lambda - t} = \left(1 - \frac{t}{\lambda}\right)^{-1}$$

$$\therefore M_X(t) = \frac{1}{2}\left(M_Y(t) + M_Y(-t)\right) = \frac{1}{2}\left(\left(1-\frac{t}{\lambda}\right)^{-1} + \left(1+\frac{t}{\lambda}\right)^{-1}\right)$$
$$(-\lambda < t < \lambda)$$

$\frac{f'}{g} + f \cdot \frac{-g'}{g^2}$

No.
Date

$\left(若\lambda = 1 \ldots \ M_X(t) = \frac{1}{2}\left(\frac{1}{1-t} + \frac{1}{1+t}\right) = \frac{1}{1-t^2} \right)$

(b) $U, V \overset{iid}{\sim} \exp(1)$

$$E[e^{t(U-V)}] = E[e^{tU}]E[e^{-tV}]$$

$$= (1-t)^{-1} \cdot (1+t)^{-1} = \frac{1}{1-t^2}$$

由此可知, $U - V \sim \text{Laplace}(\lambda = 1, \mu = 0)$

(c) $X_1 \ldots X_n \overset{iid}{\sim} \text{Laplace}(\lambda)$

$$M_X(t) = \frac{1}{2}\left\{\frac{1}{1-\frac{t}{\lambda}} + \frac{1}{1+\frac{t}{\lambda}}\right\} = \frac{1}{2} \cdot \frac{2}{1-\frac{t^2}{\lambda^2}} = \frac{\lambda^2}{\lambda^2 - t^2}$$

$$K_X(t) = \log M_X(t) = -\log(\lambda^2 - t^2)$$

$$\frac{\partial k}{\partial t} = \frac{2t}{\lambda^2 - t^2}$$

$$\frac{\partial^2 k}{\partial t^2} = \frac{2(\lambda^2 - t^2) + 4t^2}{(\lambda^2 - t^2)^2} = \frac{2\lambda^2 + 2t^2}{(\lambda^2 - t^2)^2}$$

$$\left.\frac{\partial k}{\partial t}\right|_{t=0} = 0, \quad \left.\frac{\partial^2 k}{\partial t^2}\right|_{t=0} = \frac{2\lambda^2}{\lambda^4} = \frac{2}{\lambda^2}$$

$$\therefore E[X_j] = 0 \quad V[X_j] = \frac{2}{\lambda^2}$$

$$\therefore W_n = \frac{\bar{X}}{\sqrt{\frac{1}{n} \cdot \frac{2}{\lambda^2}}} = \frac{X_1 + \ldots + X_n}{\sqrt{\frac{2}{\lambda^2} \cdot n}}$$

27

No.
Date

$$E[e^{tW_n}] = E\left[\exp\left(\frac{t\lambda}{\sqrt{2n}}(X_1+\cdots+X_n)\right)\right]$$

$$= E\left[\exp\left(\frac{t\lambda}{\sqrt{2n}}X_j\right)\right]^n$$

$$= \left(\frac{\lambda^2}{\lambda^2 - \frac{t^2\lambda^2}{2n}}\right)^n = \left(1-\frac{t^2}{2n}\right)^{-n}$$

$$\therefore\ M_{W_n}(t) = \left(1-\frac{t^2}{2n}\right)^{-n}$$

(d) $$\lim_{n\to\infty} M_{W_n}(t) = \lim_{n\to\infty}\left(1-\frac{t^2}{2n}\right)^{-n}$$

$$= \lim_{n\to\infty}\left(1-\frac{t^2}{2n}\right)^{\left(\frac{-2n}{t^2}\right)\cdot\frac{t^2}{2}}$$

$$= \exp\left(\frac{t^2}{2}\right)$$ 根據 Levy 連續性定理

$$\therefore\ W_n \xrightarrow{d} N(0,1)$$

$\boxed{13}$

(a) $X_1 \dots X_n \sim N(\mu, 1)$

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-\mu)^2\right)$$

$$f(x_1 \dots x_n \mid \mu) = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left(\sum_{j=1}^{n} -\frac{1}{2}(x_j-\mu)^2\right)$$

$$\log f(x_1 \dots x_n \mid \mu) = -\left(\frac{n}{2}\right)\log(2\pi) - \frac{1}{2}\sum_{j=1}^{n}(x_j-\mu)^2$$

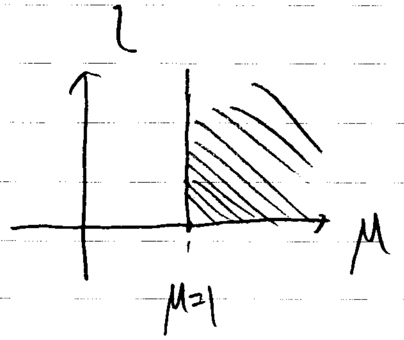
$$\frac{\partial}{\partial \mu}\log f(x_1 \dots x_n \mid \mu) = \sum_{j=1}^{n}(x_j-\mu) = 0$$

$$\Rightarrow \mu = \bar{X}$$

$$L(\mu \mid X_1 \dots X_n) = -\frac{1}{2}\sum_{j=1}^{n}(x_j-\mu)^2 - \frac{n}{2}\log 2\pi$$

$$l(\mu) = \log L$$

μ		$\bar{X}$	
l'	$+$		$-$
l	↗	MAX	↘



if .... $\bar{X} < 1 \Rightarrow \mu = 1$ 使得 L 為最大

else $\bar{X} \geq 1 \Rightarrow \mu = \bar{X}$ 使得 L 為最大

㉙

No.
Date

$\therefore\ \hat{\mu}_{MLE} = \max\{1, \bar{X}\}$

(b)(c) $\hat{\mu} = \max\{1, \bar{X}\}$ （題目是不是應改為 $\underline{\mu \geq 1}$?）

$$
\begin{aligned}
&\Pr\left(\sqrt{n}(\hat{\mu}-\mu) \leq \lambda\right) \\
&= \Pr\left(\hat{\mu}-\mu \leq \frac{\lambda}{\sqrt{n}}\right) \\
&= \Pr\left(\hat{\mu} \leq \frac{\lambda}{\sqrt{n}}+\mu\right) \\
&= \Pr\left(\max\{1, \bar{X}\} \leq \frac{\lambda}{\sqrt{n}}+\mu\right) \\
&= \Pr\left(1 \leq \frac{\lambda}{\sqrt{n}}+\mu\right) \cdot \Pr\left(\bar{X} \leq \frac{\lambda}{\sqrt{n}}+\mu\right) \\
&= \Pr\left(1 \leq \frac{\lambda}{\sqrt{n}}+\mu\right) \cdot \Pr\left(\sqrt{n}(\bar{X}-\mu) \leq \lambda\right) \\
&= \Pr\left(1 \leq \frac{\lambda}{\sqrt{n}}+\mu\right) \cdot \Phi(\lambda) \\
&= \Pr\left(\sqrt{n}(1-\mu) \leq \lambda\right) \cdot \Phi(\lambda) \\
&= I\left(\lambda \geq \sqrt{n}(1-\mu)\right) \Phi(\lambda) \quad \left(= I_{[\sqrt{n}(1-\mu), \infty)}(\lambda) \cdot \Phi(\lambda)\right)
\end{aligned}
$$

• $\mu < 1$ … $\lim\limits_{n\to\infty} \Pr\left(\sqrt{n}(\hat{\mu}-\mu) \leq \lambda\right) = 0$ (for all $-\infty < \lambda < \infty$)

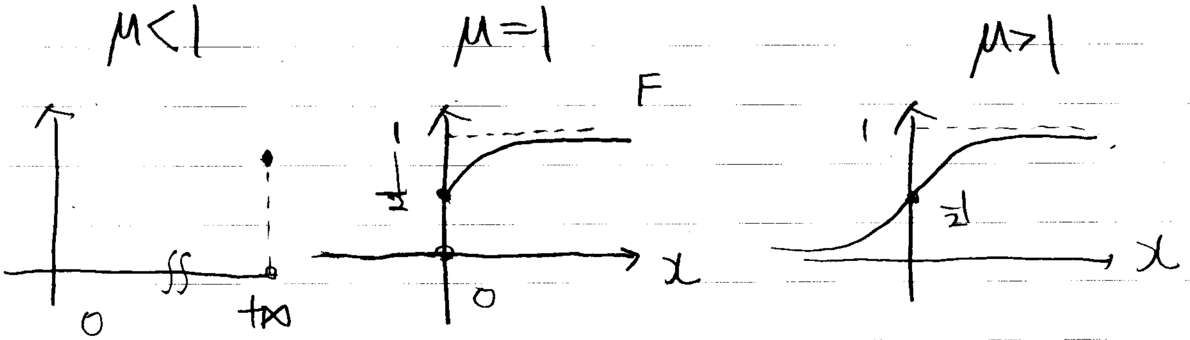
$\left(\because \lim\limits_{n\to\infty} \Pr\left(\sqrt{n}(\hat{\mu}-\mu) = +\infty\right) = 1\right)$

No.

Date

- $\mu = 1 \cdots \lim_{n \to \infty} \Pr\left(\sqrt{n}(\hat{\mu} - \mu) \leq x\right) = \begin{cases} \Phi(x) & (x \geq 0) \\ 0 & (x < 0) \end{cases}$

- $\mu > 1 \cdots \lim_{n \to \infty} \Pr\left(\sqrt{n}(\hat{\mu} - \mu) \leq x\right) = \Phi(x)$ (for all x)



③

$\boxed{14}$ $U, V \sim \exp(\lambda)$

(a) $\begin{cases} X = U+V \\ Y = UV \end{cases}$

考慮二次方程式

$$g(t) = (t-U)(t-V) = 0$$

$$\Leftrightarrow t^2 - (U+V)t + UV = 0$$

$$\Leftrightarrow t^2 - Xt + Y = 0$$

$$t = \frac{X \pm \sqrt{X^2-4Y}}{2}$$

$$\therefore \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} \frac{X+\sqrt{X^2-4Y}}{2} \\ \frac{X-\sqrt{X^2-4Y}}{2} \end{pmatrix} \text{ or } \begin{pmatrix} \frac{X-\sqrt{X^2-4Y}}{2} \\ \frac{X+\sqrt{X^2-4Y}}{2} \end{pmatrix}$$

$$J = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ v & u \end{pmatrix}$$

$$|J| = |u-v| = \sqrt{X^2-4Y}.$$

$$\frac{dxdy}{\sqrt{X^2-4Y}} = dudv$$

Chuyu culture

No.
Date

$$1=\int_{u,v\geq 0} f_{UV}(u,v)\,du\,dv = \int_{u,v\geq 0} \lambda\exp(-\lambda u)\,\lambda\exp(-\lambda v)\,du\,dv$$

$$=\int_{\substack{x^2-4y\geq 0\\ y\geq 0}} \lambda^2\exp(-\lambda x)\cdot\frac{dx\,dy}{\sqrt{x^2-4y}}\times 2$$

$\because \begin{pmatrix}u\\v\end{pmatrix}=\begin{pmatrix}\frac{x+\sqrt{x^2-4y}}{2}\\ \frac{x-\sqrt{x^2-4y}}{2}\end{pmatrix}$ or $\begin{pmatrix}\frac{x-\sqrt{x^2-4y}}{2}\\ \frac{x+\sqrt{x^2-4y}}{2}\end{pmatrix}$

$$=\int_{\substack{x^2-4y\geq 0\\ x,y>0}} \frac{2\lambda^2}{\sqrt{x^2-4y}}\exp(-\lambda x)\,dx\,dy$$

$\downarrow$

$g(t)=0$ 有兩個正實數解

$g(0)>0$, $\frac{x}{2}>0$, 判別式 ≥ 0

$$\therefore f_{XY}(x,y)=\frac{2\lambda^2}{\sqrt{x^2-4y}}\exp(-\lambda x)$$

$(x>0,\ y>0,\ x^2-4y\geq 0)$

㉝

(b) 求α,β, 使得$E[(Y-\alpha X-\beta)^2]$為最小

$$\underset{\begin{pmatrix}\alpha\\\beta\end{pmatrix}}{\parallel}\ S(\alpha,\beta) \qquad \begin{pmatrix}\alpha\\\beta\end{pmatrix} = \underset{\begin{pmatrix}\alpha\\\beta\end{pmatrix}}{\arg\min}\, E[(Y-\alpha X-\beta)^2]$$

$$\frac{\partial S}{\partial\alpha}=0 \quad \frac{\partial S}{\partial\beta}=0$$

$$\begin{cases}\alpha = E[Y] - \dfrac{Cov[X,Y]}{V[X]}E[X] \\ \beta = \dfrac{Cov(X,Y)}{V[X]}\end{cases}$$

求 X之邊際分佈…

$$\int_{0\leq y\leq\frac{x^2}{4}} \frac{2\lambda^2}{\sqrt{x^2-4y}}\exp(-\lambda x)\,dy$$

$$= 2\lambda^2\exp(-\lambda x)\cdot\left[\frac{-(x^2-4y)^{\frac{1}{2}}}{2}\right]_{y=0}^{y=\frac{x^2}{4}}$$

$$= \lambda^2 x\exp(-\lambda x)$$

$$E[X] = \int_0^\infty \lambda^2 x^2\exp(-\lambda x)\,dx = \frac{2}{\lambda}$$

$$E[X^2] = \int_0^\infty \lambda^2 x^3\exp(-\lambda x)\,dx = \frac{6}{\lambda^2}$$

$$\therefore V(X) = \frac{2}{\lambda^2}$$

No.

Date

$$E[Y] = \int_{x=0}^{x=\infty}\int_{y=0}^{y=\frac{x^2}{4}} 2\lambda^2 \exp(-\lambda x) \cdot \frac{y}{\sqrt{x^2-4y}}\, dy\, dx$$

$$= \int_{x=0}^{x=\infty}\int_{y=0}^{y=\frac{x^2}{4}} \lambda^2 \exp(-\lambda x) \frac{y}{\sqrt{\frac{x^2}{4}-y}}\, dy\, dx$$

$$= \int_{x=0}^{x=\infty}\int_{y=0}^{y=\frac{x^2}{4}} \lambda^2 \exp(-\lambda x) \cdot \frac{\left(y-\frac{x^2}{4}\right)+\left(\frac{x^2}{4}\right)}{\sqrt{\frac{x^2}{4}-y}}\, dy\, dx$$

$$= \int_{x=0}^{x=\infty}\int_{y=0}^{y=\frac{x^2}{4}} \lambda^2 \exp(-\lambda x) \left\{ \underbrace{\frac{\frac{x^2}{4}}{\sqrt{\frac{x^2}{4}-y}}}_{①} - \underbrace{\sqrt{\frac{x^2}{4}-y}}_{②} \right\} dy\, dx$$

$$① \cdots \int_{x=0}^{x=\infty}\int_{y=0}^{y=\frac{x^2}{4}} 2\lambda^2 \exp(-\lambda x) \frac{1}{\sqrt{x^2-4y}} \cdot \frac{x^2}{4}\, dx$$

$$= E\left[\frac{X^2}{4}\right] = \frac{3}{2\lambda^2}$$

$$② \cdots \int_{x=0}^{x=\infty}\int_{y=0}^{y=\frac{x^2}{4}} \lambda^2 \exp(-\lambda x) \sqrt{\frac{x^2}{4}-y}\, dy\, dx$$

$$= \int_{x=0}^{x=\infty} \left[-\lambda^2 \exp(-\lambda x)\, \frac{2}{3}\left(\frac{x^2}{4}-y\right)^{\frac{3}{2}} \right]_{y=0}^{y=\frac{x^2}{4}} dx$$

$$= \int_{x=0}^{x=\infty} \frac{2\lambda^2}{3} \cdot \frac{x^3}{8} \exp(-\lambda x)\, dx = \frac{1}{12} E(X^3) = \frac{1}{2\lambda^2}$$

$$\therefore ① - ② = \frac{1}{\lambda^2} \qquad \therefore E[Y] = \frac{1}{\lambda^2}$$

Chiyo culture

$$E[XY] = \int_{x=0}^{x=\infty}\int_{y=0}^{y=\frac{x^2}{4}} 2\lambda^2 \exp(-\lambda x)\cdot\frac{xy}{\sqrt{x^2-4y}}\,dy\,dx$$

$$\lambda^2 \exp(-\lambda x)\left\{\underbrace{\frac{\frac{x^3}{4}}{\sqrt{\frac{x^2}{4}-y}}}_{①} - \underbrace{x\sqrt{\frac{x^2}{4}-y}}_{②}\right\}$$

①… $E\left[\frac{X^3}{4}\right] = \frac{1}{4}\int_0^\infty \lambda^2 x^4 \exp(-\lambda x)\,dx$

$$= \frac{1}{4}\int_{z=0}^{z=\infty} \frac{z^4}{\lambda^2}\exp(-z)\cdot\frac{dz}{\lambda}$$

$$= \frac{\Gamma(5)}{4\lambda^3} = \frac{6}{\lambda^3}$$

②… $E\left[\frac{X^3}{12}\right] = \frac{2}{\lambda^3}$ (∵出現跟E(Y)時一樣的算式)

$① - ② = \frac{4}{\lambda^3}$

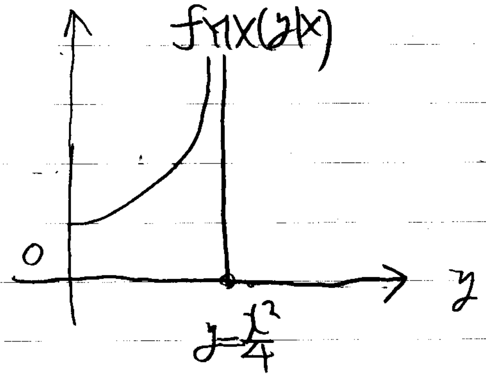
$$\mathrm{Cov}[X,Y] = E(XY) - E(X)E(Y) = \frac{4}{\lambda^3} - \frac{2}{\lambda}\frac{1}{\lambda^2} = \frac{2}{\lambda^3}$$

$$\begin{cases} \alpha = E(Y) - \dfrac{\mathrm{Cov}(X,Y)}{V(X)}E(X) = \dfrac{1}{\lambda^2} - \dfrac{\left(\frac{2}{\lambda^3}\right)}{\left(\frac{2}{\lambda^2}\right)}\cdot\dfrac{2}{\lambda} = \dfrac{1}{\lambda^2} - \dfrac{2}{\lambda^2} = -\dfrac{1}{\lambda^2} \\ \beta = \dfrac{\mathrm{Cov}[X,Y]}{V(X)} = \dfrac{\left(\frac{2}{\lambda^3}\right)}{\left(\frac{2}{\lambda^2}\right)} = \dfrac{1}{\lambda} \end{cases}$$

(c) 求 $Y|X=x$ 之分布...

$$f_{Y|X}(y|x)=\frac{f_{X,Y}(x,y)}{f_X(x)}=\frac{\frac{2\lambda^2}{\sqrt{x^2-4y}}\exp(-\lambda x)}{\lambda^2 x\exp(-\lambda x)}=\frac{2}{\sqrt{x^2-4y}\,x}$$

$$0\leq y<\frac{x^2}{4}$$



y 離 $\frac{x^2}{4}$ 越近，其機率密度越高.

因此 $\frac{x^2}{4}$ 為給定 x 之情況下的 Best Prediction

32

No.
Date

$\boxed{115}$

(a) 求 $Y|X=x$ 之分佈

X 之邊際分佈: $\int_{y=x}^{y=1} 2\,dy = 2-2x$ $(0<x<1)$

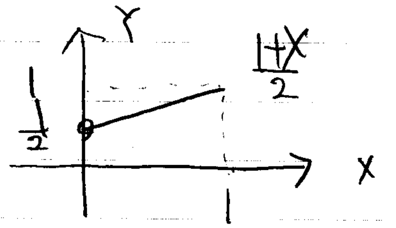
$f_X(x) = 2(1-x)$ $(\sim Be(1,2))$

$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{1}{1-x}$ $(0<x<y<1)$

由此可知 $Y|X=x \sim Uni(x,1)$

$\therefore E[Y|X=x_0] = \frac{1+x_0}{2}$

(b) $E[Y|X=x] = \frac{1+x}{2}$



$Y|X=x \sim Uni(x,1)$

均勻分佈的期望值為 $\frac{\alpha+\beta}{2}$
$(Uni(\alpha,\beta))$

因此可知 $E[Y|X=x] = \frac{1+x}{2}$

No
Date

(c) 給定 $X=x_0$ 之後，$YX^3+1 = x_0^3Y+1$

$$E[Y|X=x_0] = \frac{1+x_0}{2}$$

$$E[x_0^3Y+1|X=x_0] = \frac{x_0^4+x_0^3}{2}+1$$

(d) 考慮 Uni$(0,\theta)$ 之變異數 $(W\sim U(0,\theta))$

$$E[W] = \frac{\theta}{2}$$

$$E[W^2] = \frac{\theta^2}{3}$$

$$\therefore V[W] = \frac{\theta^2}{12}$$

$$Y|X=x_0 \sim \text{Uni}(x_0,1)$$

$V[Y|X=x_0]$ 相同於 Uni$(0,1-x_0)$ 之變異數

$$\therefore \frac{1}{12}(1-x_0)^2 \qquad (\because V[X-a]=V[X] \quad a:\text{常數})$$

給定 $X=x_0$ 之後，Y 服從 Uni$(x_0,1)$，
換言之，x_0 為 Y 之參數，所以期望值，變異 都會受其影響。

39

No.
Date

16

(a) $\frac{dF}{dx} = \frac{\beta x^{\beta-1}}{\alpha^\beta} \exp\left(-\left(\frac{x}{\alpha}\right)^\beta\right) \quad (x \geq 0)$

$f(x) = \frac{\beta x^{\beta-1}}{\alpha^\beta} \exp\left(-\left(\frac{x}{\alpha}\right)^\beta\right) \quad (x \geq 0)$

(b) $\Pr(X \leq x) = \Pr\left(\left(\frac{W}{\alpha}\right)^\beta \leq x\right)$

$= \Pr\left(\frac{W}{\alpha} \leq x^{\frac{1}{\beta}}\right) = \Pr\left(W \leq \alpha \cdot x^{\frac{1}{\beta}}\right)$

$= F\left(\alpha x^{\frac{1}{\beta}}\right) = 1 - \exp\left(-\left(\frac{\alpha x^{\frac{1}{\beta}}}{\alpha}\right)^\beta\right) = 1 - \exp(-x)$

$\therefore \quad X \sim \exp(1)$

(c) $U_1, U_2, \dots U_n \sim U(0,1)$

$X_n \overset{\text{def}}{=} n \cdot \min\{U_1, U_2, \dots, U_n\}$

$\Pr(X_n \leq x) = 1 - \Pr(X_n > x) = 1 - \Pr(n \cdot \min\{U_1 \dots U_n\} > x)$

$= 1 - \Pr\left(\min\{U_1 \dots U_n\} > \frac{x}{n}\right)$

$= 1 - \left(1 - \frac{x}{n}\right)^n = 1 - \left(1 - \frac{x}{n}\right)^{\frac{-n}{x} \cdot (-x)}$

$\lim_{n \to \infty} \Pr(X_n \leq x) = 1 - e^{-x}$

$\therefore \quad X_n \xrightarrow{d} \exp(1)$

No.
Date

$\therefore X_n = \left(\frac{W_n}{\alpha}\right)^{\beta}$ （利用(b)的結果）

$W_n = \alpha X_n^{\frac{1}{\beta}}$

$W_n \xrightarrow{d} \text{Weibull}(\alpha, \beta)$ $\therefore W_n = \alpha\left(n \cdot \min\{X_1 \cdots X_n\}\right)^{\frac{1}{\beta}}$

$\boxed{17}$ (a) $X_1, X_2 \sim \text{Uni}(0, \theta)$

- $f_{X_1 X_2}(x_1, x_2) = \frac{1}{\theta^2} I_{(0,\theta)}(x_1)\, I_{(0,\theta)}(x_2)$
- $\Pr(T \le t) = \Pr(X_1, X_2 \le t) = \frac{t^2}{\theta^2}$

$\frac{d}{dt}\Pr(T \le t) = \frac{2t}{\theta^2} \quad (0 < t < \theta)$

$f_T(t) = 2t\, I_{(0,\theta)}(t) / \theta^2$

- $f_{X_1, X_2, t}(x_1, x_2, t) = \begin{cases} \text{if } X_1 = t \cdots & \frac{1}{\theta^2} I_{(0,t)}(x_2)\, I_{(0,\theta)}(t) \\ \text{else } X_2 = t & \frac{1}{\theta^2} I_{(0,t)}(x_1)\, I_{(0,\theta)}(t) \end{cases}$

$= \frac{1}{\theta^2} \cdot I_{(0,t)}(x_2)\, I_{(0,\theta)}(t)\, I(t = x_1) + \frac{1}{\theta^2} I_{(0,t)}(x_1)\, I_{(0,\theta)}(t)\, I(t = x_2)$

$\therefore \frac{f_{X_1, X_2, t}(x_1, x_2, t)}{f_T(t)} = \frac{1}{2t}\left\{ I_{(0,t)}(x_2)\, I(t = x_1) + I_{(0,t)}(x_1)\, I(t = x_2) \right\}$

No.
Date

(b) 求 θ 之最小充份統計量，

$\frac{f(Y_1=y_1, Y_2=y_2)}{f(X_1=x_1, X_2=x_2)}$ 與 θ 無關 $\Rightarrow$

$$\frac{\frac{1}{\theta^2} I(0<Y_{(1)}) I(Y_{(2)}<\theta)}{\frac{1}{\theta^2} I(0<X_{(1)}) I(X_{(2)}<\theta)} = \frac{I(0<Y_{(1)})}{I(0<X_{(1)})} \cdot \frac{I(Y_{(2)}<\theta)}{I(X_{(2)}<\theta)}$$

與 θ 無關 $\Rightarrow$ $X_{(2)}=Y_{(2)}$ 為必要條件.

若 $X_{(2)}=Y_{(2)} \Rightarrow \frac{f(Y_1=y_1, Y_2=y_2)}{f(X_1=x_1, X_2=x_2)}$ 與 θ 無關.

$\therefore$「$X_{(2)}=Y_{(2)}$」為 $\frac{f(Y_1=y_1, Y_2=y_2)}{f(X_1=x_1, X_2=x_2)}$ 與 θ 無關

之必要充份條件.

故此，$X_{(2)}=\max\{X_1, X_2\}$ 為 θ 之最小充份統計量

我們無法由 X_1+X_2 得知 $\max\{X_1, X_2\}$ 之值，

$\therefore$ X_1+X_2 並非 θ 之充份統計量。

$\boxed{18}$. $f(x|\theta) = \frac{1}{3\theta} I(-\theta < x < 2\theta)$

$$f(x_1 \ldots x_n|\theta) = \left(\frac{1}{3\theta}\right)^n I(-\theta < x_1 < 2\theta) \cdots I(-\theta < x_n < 2\theta)$$

$$= \left(\frac{1}{3\theta}\right)^n I(-\theta < X_{(1)},\ X_{(n)} < 2\theta)$$

$$= \left(\frac{1}{3\theta}\right)^n I\left(-X_{(1)} < \theta \cap \frac{X_{(n)}}{2} < \theta\right)$$

$$= \left(\frac{1}{3\theta}\right)^n I\left(\theta > \max\left\{-X_{(1)}, \frac{X_{(n)}}{2}\right\}\right)$$

$$= \left(\frac{1}{3\theta}\right)^n I_{\left(\max\left\{-X_{(1)}, \frac{X_{(n)}}{2}\right\}, \infty\right)}(\theta)$$

$$\therefore L(\theta|X_1 \ldots X_n) = \left(\frac{1}{3\theta}\right)^n I_{\left(\max\left\{-X_{(1)}, \frac{X_{(n)}}{2}\right\}, \infty\right)}(\theta)$$

(由此可知, $\hat{\theta}_{MLE} = \max\left\{-X_{(1)}, \frac{X_{(n)}}{2}\right\}$)

Date

$\boxed{19}$

$$\hat{f}_n(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x-x_i}{h}\right)$$ (Kernel Density Estimation)

$n=4,\ h=1.$

$K=\phi$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0.5 \\ 3 \end{pmatrix}$$

$$\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \phi(x-x_i)$$

① $x=0 \cdots$ $\hat{f}_n(0) = \frac{1}{n}\sum_{i=1}^{n}\phi(-x_i) = \frac{1}{4}\left(\phi(1)+\phi(0)+\phi(-0.5)+\phi(-3)\right)$

② $x=0.5 \cdots$ $\hat{f}_n(0.5) = \frac{1}{n}\sum_{i=1}^{n}\phi(0.5-x_i) = \frac{1}{4}\left(\phi(1.5)+\phi(0.5)+\phi(0)+\phi(-2.5)\right)$

③ $x=1 \cdots$ $\hat{f}_n(1) = \frac{1}{n}\sum_{i=1}^{n}\phi(1-x_i) = \frac{1}{4}\left(\phi(2)+\phi(1)+\phi(0.5)+\phi(-2)\right)$

$$\rightarrow \begin{cases} ① = 0.249353 \\ ③ = 0.224513 \\ ③ = 0.175504 \end{cases}$$

$\boxed{20}$

(a) $\hat{f}_n(x) \mid X_1, X_2 \dots X_n = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{x-x_i}{h}\right)$

(b) $X \mid X_1 \dots X_n \sim \hat{f}_n(x)$

$$E[X \mid X_1, \dots, X_n] = \int_{x\in\mathbb{R}} \frac{x}{nh}\sum_{j=1}^{n} K\left(\frac{x-x_i}{h}\right) dx$$

$$= \frac{1}{nh}\sum_{j=1}^{n} \underbrace{\int_{x\in\mathbb{R}} x K\left(\frac{x-x_i}{h}\right) dx}$$

$$y = \frac{x-x_i}{h} \qquad \frac{dy}{dx} = \frac{1}{h}$$

$$= \frac{1}{nh}\sum_{j=1}^{n} \int_{y\in\mathbb{R}} (hy + x_i) K(y)\, h\, dy$$

$$= \frac{1}{nh}\sum_{j=1}^{n} h x_i = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{X_1 + \dots + X_n}{n} = \bar{X}$$

(c) $$E[X^2 \mid X_1 \dots X_n] = \int_{x\in\mathbb{R}} \frac{x^2}{nh}\sum_{j=1}^{n} K\left(\frac{x-x_i}{h}\right) dx$$

$$= \frac{1}{nh}\sum_{j=1}^{n} \int_{y\in\mathbb{R}} (hy + x_i)^2 K(y) \cdot h\, dy$$

$$= \frac{1}{n}\sum_{j=1}^{n} \int_{y\in\mathbb{R}} (h^2y^2 + 2hx_i y + x_i^2) K(y)\, dy$$

$$= \frac{1}{n}\sum_{j=1}^{n} (h^2\sigma_K^2 + x_i^2) = h^2\sigma_K^2 + \frac{1}{n}\sum_{j=1}^{n} x_i^2$$

Date

$$V[X|X_1,X_2\ldots X_n] = h^2\sigma_K^2 + \frac{1}{n}\sum_{j=1}^{n} x_i^2 - \bar{X}^2$$
$$= h^2\sigma_K^2 + \frac{1}{n}\sum_{j=1}^{n}(x_j-\bar{x})^2$$

(d) $E[X|X_1=x_1\ldots X_n=x_n) = \bar{X}$

$$V[X|X_1=x_1\ldots X_n=x_n) = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 + h^2\sigma_K^2$$

現在假設 $X_1,X_2\ldots X_n$ 存在 $E[X_j]$ 及 $E[X_j^2]$

(換言之 $|E[X_j]|<\infty$, $E[X_j^2]<\infty$)

、根據弱大數法則, $\bar{X}\xrightarrow{P}\mu$ (X_j之期望值)

(Khinchine 的弱大數法則) $\frac{1}{n}\sum_{j=1}^{n}(X_j-\bar{X})^2\xrightarrow{P}\sigma^2$ (X_j之變異數)

、所以 $n\to\infty$ 時, $h^2\sigma_K^2$ 需趨近於 0,

(註) Khinchine 弱大數法則

$X_1,X_2\ldots X_n$ 之 $|E[X_j]|<\infty \Rightarrow \frac{X_1+\ldots+X_n}{n}\xrightarrow{P} E[X_1]=\mu$

$$M_{\bar{X}}(t) = E[e^{t\bar{X}}] = E[e^{\frac{t}{n}X_j}]^n = M_X(\tfrac{t}{n})^n$$

泰勒展開 $M_X(t) = M(0)+M'(0)\,t+\frac{M''(0)}{2!}t^2+\ldots$

$$M_X(\tfrac{t}{n}) = 1+\mu\cdot\frac{t}{n}+o(\tfrac{1}{n^2})$$

$n\to\infty$ $M_X(\frac{t}{n})^n \longrightarrow \lim_{n\to\infty}(1+\frac{\mu t}{n})^{\frac{n}{\mu t}\cdot\mu t} = e^{\mu t}$ $\therefore \bar{X}\xrightarrow{P}\mu$

No.
Date

$\boxed{21}$

$(y>0)$

(a) $\Pr(Y\leqq y)=\Pr(-\log X\leqq y)=\Pr(\exp(-\log X)\leqq e^{y})$

$=\Pr(\frac{1}{X}\leqq e^{y})=\Pr(e^{-y}\leqq X)=1-\Pr(X<e^{-y})$

$\left(\because \Pr(X\leqq \lambda)\ (0<\lambda<1)=\int_0^{\lambda}\theta t^{\theta-1}dt = [t^{\theta}]_0^{\lambda}=\lambda^{\theta}\right)$

$\therefore \Pr(Y\leqq y)=1-\Pr(X<e^{-y})=1-(e^{-y})^{\theta}=1-e^{-\theta y}$

由此可知, $Y\sim \exp(\theta)$ (mean $\frac{1}{\theta}$ 之指數分布)

(b) 假設T為 $\frac{1}{\theta}$ 之不偏估計量

$V[T]\geqq \frac{(\tau'(\theta))^2}{I_n(\theta)}=\frac{(\tau'(\theta))^2}{nI_1(\theta)}$

$\tau'(\theta)=-\frac{1}{\theta^2}\quad (\tau'(\theta))^2=\frac{1}{\theta^4}$

觀測到一個 X_i 時, 其 Fisher 資訊量 $I_1(\theta)=\frac{1}{\theta^2}$

$\left(\because \frac{\partial}{\partial\theta}\log f(x|\theta)=\frac{\partial}{\partial\theta}\{(\theta-1)\log X+\log\theta\}=\log X+\frac{1}{\theta}\right.$

$\left.E\left[\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right)^2\right]=E\left[\left(-\log X-\frac{1}{\theta}\right)^2\right]=V[-\log X]=\frac{1}{\theta^2}\right.$

(41)

No.
Date

[2] (b)

$I_1(\theta) = \frac{1}{\theta^2}$ $I_n(\theta) = \frac{n}{\theta^2}$

$\therefore V[\,] \geq \frac{(\tau'(\theta))^2}{I_n(\theta)} = \frac{\frac{1}{\theta^4}}{\left(\frac{n}{\theta^2}\right)} = \frac{1}{n\theta^2}$ である

故, Cramer Rao's Lower Bound $= \frac{1}{n\theta^2}$

(c) 方法1 證明 $\sum_{j=1}^{n} \frac{-\log X_j}{n}$ 為 $\frac{1}{\theta}$ 之有效估計量,

- $-\log X_j \sim \exp(\theta)$ (mean: $\frac{1}{\theta}$, var: $\frac{1}{\theta^2}$)
- $\sum_{j=1}^{n} -\log X_j \sim \Gamma(n, \theta)$ (mean $\frac{n}{\theta}$, var $\frac{n}{\theta^2}$)
- $\frac{1}{n}\sum_{j=1}^{n} -\log X_j$ 期望值 $\frac{1}{\theta}$, 變異數 $\frac{1}{n\theta^2}$ (滿足 Cramer Rao's Lower Bound)

由此可知, $\frac{1}{n}\sum_{j=1}^{n} -\log X_j$ 為 $\frac{1}{\theta}$ 之不偏估計量,

且達到 Cramer Rao's Lower Bound, 所以此估計量

為有效估計量, 有效估計量必定為 UMVUE.

Chuyu culture

方法2 $f(x_1,\ldots,x_n|\theta)=\theta^n(x_1x_2\cdots x_n)^{\theta-1}I_{(0,1)}(x_1,\ldots,x_n)$

$=\theta^n(x_1x_2\cdots x_n)^{-1}\cdot\exp\left(\theta\sum_{i=1}^{n}\log x_i\right)$

此分佈屬於指數族。

其中 $\sum_{i=1}^{n}\log x_i$ 之係數為 θ。

$\{\theta|\theta\in\Theta\}(CR)$ 顯然包含內點(interior point)

故此，$\sum_{i=1}^{n}\log x_i$ 為 θ 之完備充份統計量

如方法1所示，$-\sum_{i=1}^{n}\log x_i/n$ 為 θ 之不偏推定量

且完備統計量 $\sum_{i=1}^{n}\log x_i$ 之函數。

根據 Lehman-Scheffe 定理，$-\sum_{i=1}^{n}\frac{1}{n}\log x_i$ 為 θ 之UMVUE。

22

(a) $f(x|\theta) = \frac{1}{\theta} I(0<x<\theta)$

$f(x_1 \dots x_n|\theta) = \frac{1}{\theta^n} I(0<X_{(1)}, X_{(n)}<\theta)$

$$= \underbrace{I_{(0,\infty)}(X_{(1)})}_{h(x)} \cdot \underbrace{\frac{1}{\theta^n} I_{(-\infty,\theta)}(X_{(n)})}_{g(X_{(n)}|\theta)}$$

根據 Neyman-Fisher 的分解定理,

$X_{(n)}$ 為 θ 之充份統計量,

(a)' 為了方便起見, 在此證明 $X_{(n)}$ 為 θ 之只備充份統計量,

$Pr(X_{(n)} \le t) = Pr(X_1, X_2 \dots X_n \le t) = \left(\frac{t}{\theta}\right)^n \quad (0<t\le\theta)$

$\frac{d}{dt} Pr(X_{(n)} \le t) = \frac{nt^{n-1}}{\theta^n}$

$\therefore f_{X_{(n)}}(t) = \frac{nt^{n-1}}{\theta^n} \quad (0<t\le\theta) \dots X_{(n)}$ 之機率密度函數,

$E[g(X_{(n)})] = \int_0^\theta g(t) \cdot \frac{nt^{n-1}}{\theta^n} dt = 0 \qquad (\times \frac{\theta^n}{n})$

$\Rightarrow \int_0^\theta g(t) t^{n-1} dt = 0$

$\Rightarrow \frac{d}{d\theta} \int_0^\theta g(t) t^{n-1} dt = g(\theta) \cdot \theta^{n-1} = 0$

No.
Date

$\Rightarrow g(\theta)=0$ (for all $\theta>0$)

$\therefore E[g(X_{(n)})]=0 \Rightarrow Pr(g(X_{(n)})=0)=1$

由此可知 $X_{(n)}$ 為 θ 之完備充份統計量

(b) $E[\hat{\theta}]= E[2X_1]=\int_0^\theta 2x\frac{1}{\theta}dx=\left[\frac{x^2}{\theta}\right]_0^\theta=\frac{\theta^2}{\theta}=\theta$

$\therefore \hat{\theta}$ 為 θ 之不偏估計量

(c) $\hat{\theta}$ 為由 $\begin{pmatrix}X_1\\ \vdots \\ X_n\end{pmatrix}$ 構成的統計量

$X_{(n)}$ 為充份統計量，

$\hat{\theta}\mid X_{(n)}$ 的分佈與 θ 無關，

因此，$E[\hat{\theta}\mid X_{(n)}=t]$ 為 t 之函數，

如(a)'所述，$X_{(n)}$ 為 θ 之完備充份統計量

所以 $E[\hat{\theta}\mid X_{(n)}=t]$ 為完備充份統計量 $X_{(n)}(t)$ 為函數

而且，$E[E[\hat{\theta}\mid X_{(n)}]]=E[\hat{\theta}]=\theta$，

所以 $E[\hat{\theta}\mid X_{(n)}]$ 為 θ 之不偏估計量

根據 Lehman-scheff 定理，$E[\hat{\theta}\mid X_{(n)}]$ 為 θ 之 UMVUE

我們找出一個 $X_{(n)}$ 之函數滿足 $E[g(X_{(n)})]=\theta$

$f_{X_{(n)}}(t)=\frac{nt^{n-1}}{\theta^n}\quad(0<t\leq\theta)$

$E[X_{(n)}]=\int_0^\theta \frac{nt^n}{\theta^n}dt=\frac{n}{n+1}\left[\frac{t^{n+1}}{\theta^n}\right]_0^\theta=\frac{n}{n+1}\theta$

$\therefore E\left[\frac{n+1}{n}X_{(n)}\right]=\theta$

由此可知 $E[\hat{\theta}\,|\,X_{(n)}]=\frac{n+1}{n}X_{(n)}$ 為 θ 之 UMVUE. ⌟

※ 別的方法… 直接求 $E[\hat{\theta}\,|\,X_{(n)}]$

$\Downarrow$

先考慮 $X_1\,|\,X_{(n)}=t$ 之條件分佈

$P_r(X_1\leq x_1,\ X_{(n)}\leq t)$

$=P_r(X_1\leq x_1,\ X_1\leq t,\ X_2\leq t,\cdots X_n\leq t)$

$=P_r(X_1\leq \min\{x_1,t\},\ X_2\sim X_n\leq t)$

$=\left(\frac{\min\{x_1,t\}}{\theta}\right)\cdot\frac{t^{n-1}}{\theta^{n-1}}$

$=\frac{\min\{x_1,t\}\,t^{n-1}}{\theta^n}$

No.
Date

- case 1 ... $x_1 \le t$ ... $Pr(X_1 \le x_1, X_{(n)} \le t) = \dfrac{x_1 \cdot t^{n-1}}{\theta^n}$

$$\frac{\partial^2}{\partial x_1 \partial t} Pr(X_1 \le x_1, X_{(n)} \le t) = \frac{(n-1)t^{n-2}}{\theta^n}$$

- case 2 ... $x_1 > t$ ... $Pr(X_1 \le x_1, X_{(n)} \le t) = \dfrac{t^n}{\theta^n}$

$$\frac{\partial^2}{\partial x_1 \partial t} Pr(X_1 \le x_1, X_{(n)} \le t) = 0$$

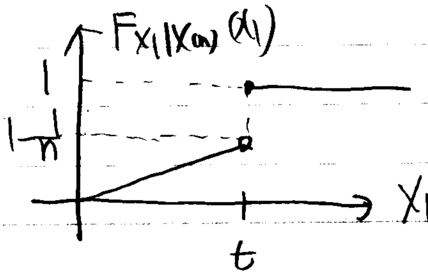
$$\therefore f_{X_1, X_{(n)}}(x_1, t) = \frac{(n-1)t^{n-2}}{\theta^n} \cdot I(x_1 < t)$$

$$f_{X_{(n)}}(t) = \frac{n t^{n-1}}{\theta^n}$$

$$\therefore f_{X_1 | X_{(n)}}(X_1 = x_1 | X_{(n)} = t) = \frac{\frac{(n-1)t^{n-2}}{\theta^n} I_{(0,t)}(x_1)}{\left(\frac{n t^{n-1}}{\theta^n}\right)}$$

$$= \frac{n-1}{n} \cdot \frac{1}{t} \cdot I_{(0,t)}(x_1)$$

由此可知 $X_1 | X_{(n)} = t$ 之條件累積分佈函數如下:



$$F_{X_1|X_{(n)}}(x_1) = \begin{cases} \frac{n-1}{n} \cdot \frac{x_1}{t} & (0 < x_1 < t) \\ 1 & (x_1 \ge t) \end{cases}$$

$(Pr(X_1 = t) = \frac{1}{n})$

No.
Date

$\therefore E[X_1 \mid X_{(n)}=t] = \int_0^t \frac{n-1}{n} \cdot \frac{x_1}{t}\, dx_1 + t \cdot \underbrace{P_r(X_1=t)}_{\frac{1}{n}}$

$= (1-\frac{1}{n}) \cdot \left[\frac{x_1^2}{2t}\right]_0^t + \frac{t}{n}$

$= (1-\frac{1}{n}) \cdot \frac{t}{2} + \frac{t}{n}$

$= (\frac{1}{2} - \frac{1}{2n} + \frac{1}{n})t = (\frac{1}{2} + \frac{1}{2n})t$

$\therefore E[\hat{\theta} \mid X_{(n)}=t] = E[2X_1 \mid X_{(n)}=t]$

$= 2(\frac{1}{2} + \frac{1}{2n})t = (1+\frac{1}{n})t$

$\therefore E[\hat{\theta} \mid X_{(n)}=t] = (\frac{n+1}{n})t$

$\therefore \theta$之UMVUE為 $\frac{n+1}{n} X_{(n)}$

(得到跟第一個做法一樣的結論)

No

Date

$\boxed{23}$

(a) $Pr(X_1=x_1) = e^{-\lambda} \cdot \dfrac{\lambda^{x_1}}{x_1!}$

$$Pr(X_1=x_1, \dots, X_n=x_n) = e^{-n\lambda} \cdot \frac{\lambda^{(x_1+\dots+x_n)}}{(x_1!\dots x_n!)}$$

$$= \underbrace{\frac{1}{x_1!\dots x_n!}}_{h(\vec{x})} \cdot \underbrace{e^{-n\lambda} \cdot \lambda^{(x_1+\dots+x_n)}}_{g(X_1+X_2+\dots+X_n \mid \lambda)}$$

根據Neyman-Fisher的分解定理，

$X_1+X_2+\dots+X_n$ 為 θ 之充份統計量，

(b) $Pr(\hat{\theta}=1) = Pr(X_1=0) = e^{-\lambda} \cdot \dfrac{\lambda^0}{0!} = e^{-\lambda}$

$$\therefore E[\hat{\theta}] = \sum_{a=0,1} a \cdot Pr(\hat{\theta}=a) = Pr(\hat{\theta}=1) = e^{-\lambda}$$

$\therefore \hat{\theta}$ 為 $e^{-\lambda}=\theta$ 之不偏估計量，

(c) $E[\hat{\theta} \mid T=t] = Pr(\hat{\theta}=1 \mid T=t) = Pr(X_1=0, T=t)$

$$= \frac{Pr(X_1=0, T=t)}{Pr(T=t)} = \frac{Pr(X_1=0, X_1+X_2+\dots+X_n=t)}{Pr(T=t)}$$

(55)

No.
Date

$$\frac{Pr(X_1=0,\ X_2+X_3+\cdots X_n=t)}{Pr(T=t)} = \frac{Pr(X_1=0)\,Pr(X_2+\cdots+X_n=t)}{Pr(X_1+X_2+\cdots X_n=t)}$$

$$\begin{cases} \bullet\ X_1+X_2+\cdots X_n \sim Po(n\lambda) \\ \bullet\ X_2+X_3+\cdots X_n \sim Po((n-1)\lambda) \end{cases}$$

$$= \frac{e^{-\lambda}\cdot e^{-(n-1)\lambda}\cdot\frac{((n-1)\lambda)^t}{t!}}{\left(e^{-n\lambda}\cdot\frac{(n\lambda)^t}{t!}\right)} = \frac{\{(n-1)\lambda\}^t}{(n\lambda)^t} = \left(1-\frac{1}{n}\right)^t$$

$$\therefore\ E[\hat{\theta}\,|\,T=t] = \left(1-\frac{1}{n}\right)^t = \left(1-\frac{1}{n}\right)^{(x_1+x_2+\cdots+x_n)}$$

No.
Date

$\boxed{24}$

(a) $\int_0^\infty x^2 \frac{1}{\theta} \exp\left(\frac{-x}{\theta}\right) dx \qquad y = \frac{x}{\theta} \quad \frac{dy}{dx} = \frac{1}{\theta}$

$= \int_0^\infty (\theta y)^2 \cdot \frac{1}{\theta} \exp(-y) \cdot \theta \, dy$

$= \int_0^\infty \theta^2 y^2 \exp(-y) dy = \theta^2 \cdot \Gamma(3) = 2\theta^2$

$\therefore E[X_i^2] = 2\theta^2$

$\therefore E[\frac{1}{2} X_i^2] = \theta^2$

$\therefore \frac{1}{2} X_i^2$ 為 θ^2 之不偏估計量

(b) $f(x|\theta) = \frac{1}{\theta} \exp\left(\frac{-x}{\theta}\right)$

$f(x_1 \cdots x_n | \theta) = \frac{1}{\theta^n} \exp\left(\frac{-1}{\theta}(x_1 + x_2 + \cdots + x_n)\right)$

$\left(\exists g \text{ s.t.}\right) \quad = g(X_1 + X_2 + \cdots X_n | \theta)$

$\therefore$ 根據 Neyman-Fisher 的分解定理

$X_1 + X_2 + \cdots + X_n$ 為 θ 之充份統計量

51

No
Date

(c) $f(x_1, x_2, \ldots x_n \mid \theta) = \frac{1}{\theta^n} \exp\left(-\frac{1}{\theta}(x_1+x_2+\cdots+x_n)\right)$

exp内 $(X_1+X_2+\cdots X_n)$ 之係數為 $-\frac{1}{\theta}$

$\left\{-\frac{1}{\theta} \mid \theta>0\right\}$ 可内含 R上的開集合 G.

$\therefore$ $X_1+X_2+\cdots X_n$ 為 θ 之完備充份統計量.

根據 Rao-Blackwell & Lehman scheffe 定理

$E\left[\frac{X_1^2}{2} \mid X_1+X_2+\cdots X_n = t\right]$ 為 θ^2 之 UMVUE.

① 方法1: 我們知道 $E\left[\frac{X_1^2}{2} \mid X_1+X_2+\cdots+X_n = t\right)$ 為

T之函數, 而且, θ^2 之不偏估計量

$T = X_1+X_2+\cdots+X_n \sim \Gamma(n, \frac{1}{\theta})$

$E[T^2] = \int_0^\infty \frac{t^{n+1}}{\Gamma(n)\cdot\theta^n} \exp\left(-\frac{t}{\theta}\right) dt$ $\left(u = \frac{t}{\theta}, \frac{du}{dt} = \frac{1}{\theta}\right.$

$= \int_0^\infty \frac{(\theta u)^{n+1}}{\Gamma(n)\,\theta^n} \exp(-u)\cdot\theta\, du$

$= \theta^2 \cdot \frac{\Gamma(n+2)}{\Gamma(n)} = \theta^2 \cdot \frac{(n+1)!}{(n-1)!} = \theta^2 \cdot n\cdot(n+1)$

No

Date

$\therefore E\left[\frac{T^2}{n(n+1)}\right]=\theta^2$

$\frac{T^2}{n(n+1)}$ 為完備統計量T之函數，且滿足 $E\left[\frac{T^2}{n(n+1)}\right]=\theta^2$

$\therefore \frac{T^2}{n(n+1)}$ 為 θ^2 之UMVUE，（∵ Rao-Blackwell & Lehman scheffe 定理）

$\therefore E\left[\frac{X_1^2}{2} \middle| X_1+\cdots+X_n=T\right]$ 亦為 $\frac{T^2}{n(n+1)}$

② 方法2：直接求 $E\left(\frac{X_1^2}{2} \middle| X_1+X_2+\cdots+X_n=t\right)$

先考慮 $X_1 \mid X_1+X_2+\cdots+X_n=t$ 之分佈（但先求 X_1, T 之聯合分佈）

$$\begin{cases} W=X_1 & \sim \exp\left(\frac{1}{\theta}\right) \\ Z=X_2+X_3+\cdots+X_n & \sim \Gamma\left(n-1, \frac{1}{\theta}\right) \end{cases}$$

W, Z：獨立，$f_{W,Z}(w,z)=\frac{1}{\theta}\exp\left(\frac{-w}{\theta}\right)\cdot\frac{z^{n-2}}{\Gamma(n-1)\cdot\theta^{n-1}}\exp\left(\frac{-z}{\theta}\right)$

$(w\geq 0,\ z\geq 0)$

$$\begin{cases} X_1=W \\ T=W+Z \end{cases}$$ （變數轉換）

$\frac{\partial}{\partial w}\begin{pmatrix} x_1 \\ t \end{pmatrix}=\begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \frac{\partial}{\partial z}\begin{pmatrix} x_1 \\ t \end{pmatrix}=\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$\therefore J=\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ $\quad \det J=1$ $\quad dx_1\,dt=dw\,dz$

(59)

No
Date

$$\begin{cases} W \geq 0, \\ Z \geq 0 \end{cases} \Rightarrow \begin{cases} X_1 \geq 0 \\ T \geq X_1 \end{cases}$$

全機率 $=1=\iint_{w,z\geq 0} \frac{1}{\theta}\exp\left(\frac{-w}{\theta}\right)\cdot\frac{1}{\theta^{n-1}}\cdot\frac{z^{n-2}}{\Gamma(n-1)}\exp\left(\frac{-z}{\theta}\right)dwdz$

$=\iint_{t\geq x_1\geq 0} \frac{1}{\theta^n}\cdot\frac{1}{\Gamma(n-1)}\cdot(t-x_1)^{n-2}\cdot\exp\left(\frac{-t}{\theta}\right)\cdot dx_1 dt$

$\therefore f_{X_1,T}(x_1,t)=\frac{1}{\theta^n}\cdot\frac{1}{\Gamma(n-1)}(t-x_1)^{n-2}\exp\left(\frac{-t}{\theta}\right)\quad(0\leq x_1\leq t)$ //

(X_1, T之聯合分佈)

$T\sim\Gamma\left(n,\frac{1}{\theta}\right)\qquad \therefore f_T(t)=\frac{t^{n-1}}{\Gamma(n)\cdot\theta^n}\exp\left(\frac{-t}{\theta}\right)\quad(t\geq 0)$

故, $f_{X_1|T}(x_1|t)=\dfrac{\left\{\dfrac{(t-x_1)^{n-2}}{\Gamma(n-1)\,\theta^n}\exp\left(\dfrac{-t}{\theta}\right)\right\}}{\left\{\dfrac{t^{n-1}}{\Gamma(n)\,\theta^n}\exp\left(\dfrac{-t}{\theta}\right)\right\}}\quad\left(=\dfrac{f_{X_1,T}(x_1,t)}{f_T(t)}\right)$

$=(n-1)\cdot\frac{1}{t}\left(1-\frac{x_1}{t}\right)^{n-2}\quad(0\leq x_1\leq t)$

↓

$X_1|T=t$ 之條件分佈

No.
Date

$\therefore E\left[\frac{X_1^2}{2}\middle|T=t\right]=\int_0^t \frac{x_1^2}{2}\cdot(n-1)\cdot\frac{1}{t}\cdot\left(1-\frac{x_1}{t}\right)^{n-2}dx_1$

$=\frac{(n-1)}{2t}\int_0^t x_1^2\left(1-\frac{x_1}{t}\right)^{n-2}dx_1$

$\left(\text{令 } \frac{x_1}{t}=y \quad \frac{dy}{dx_1}=\frac{1}{t}\right)$

$=\frac{n-1}{2t}\int_0^1 (ty)^2(1-y)^{n-2}\,t\,dy$

$=\left(\frac{n-1}{2}\right)\cdot t^2\int_0^1 y^2(1-y)^{n-2}dy$

$=\left(\frac{n-1}{2}\right)t^2\cdot Be(3,n-1)$

$=\left(\frac{n-1}{2}\right)t^2\cdot\frac{\Gamma(3)\Gamma(n-1)}{\Gamma(n+2)}=\left(\frac{n-1}{2}\right)\cdot t^2\cdot\frac{2!\,(n-2)!}{(n+1)!}$

$=(n-1)t^2\cdot\frac{1}{(n+1)\cdot n\cdot(n-1)}$

$=\frac{t^2}{n(n+1)}$

$\therefore E\left[\frac{X_1^2}{2}\middle|T=t\right]=\frac{t^2}{n(n+1)}$ $\therefore \frac{(X_1+X_2+\cdots+X_n)^2}{n(n+1)}$ 為 θ^2 之 UMVUE

(得到跟方法①相同的結論)

61

No.

Date

(d) 求 θ 之 MLE,

$$f(x_1,x_2,\dots,x_n|\theta)=\frac{1}{\theta^n}\exp\left(-\frac{1}{\theta}(x_1+x_2+\dots+x_n)\right)$$

$$\log f(x_1,x_2,\dots,x_n|\theta)=-n\log\theta-\frac{1}{\theta}(x_1+x_2+\dots+x_n)$$

$$\frac{\partial}{\partial\theta}\log f(x_1,x_2,\dots,x_n|\theta)=-\frac{n}{\theta}+\frac{1}{\theta^2}(x_1+x_2+\dots+x_n)$$

$$=\frac{n\bar{X}-n\theta}{\theta^2}=\frac{n(\bar{X}-\theta)}{\theta^2}$$

θ	0		$\bar{X}$		
$\frac{\partial}{\partial\theta}\log f$		+	0	−	
f		$\nearrow$	max	$\searrow$	

$\therefore$ $\theta=\bar{X}$ 使得 $L(\theta|x_1\dots x_n)=f(x_1\dots x_n|\theta)$

為最大、 $\therefore$ $\hat{\theta}_{MLE}=\bar{X}$

根據 MLE 的不變性, θ^2 的 MLE 為 $(\hat{\theta})^2=\bar{X}^2$

比較 $MSE\left(\frac{(X_1+\dots+X_n)^2}{n(n+1)}\right)$ vs $MSE\left(\frac{(X_1+\dots+X_n)^2}{n^2}\right)$

① ②

No

Date

① $MSE\left(\frac{(X_1+\cdots X_n)^2}{n(n+1)}\right) = MSE\left(\frac{T^2}{n(n+1)}\right)$

$= E\left[\left(\frac{T^2}{n(n+1)} - \theta^2\right)^2\right] = V\left[\frac{T^2}{n(n+1)}\right]$

$= \underbrace{\frac{1}{n^2(n+1)^2}V[T^2]}$

② $MSE\left(\frac{(X_1+\cdots X_n)^2}{n^2}\right) = MSE\left(\frac{T^2}{n^2}\right)$

$= E\left[\left(\frac{T^2}{n^2} - \theta^2\right)^2\right] \qquad \left(E\left[\frac{T^2}{n^2}\right] = \frac{n+1}{n}\theta^2\right)$

$= E\left[\left(\frac{T^2}{n^2} - \frac{n+1}{n}\theta^2 + \frac{\theta^2}{n}\right)^2\right]$

$= E\left[\left(\frac{T^2}{n^2} - \frac{n+1}{n}\theta^2\right)^2 + 2\left(\frac{T^2}{n^2} - \frac{n+1}{n}\theta^2\right)\cdot\frac{\theta^2}{n} + \frac{\theta^4}{n^2}\right]$

$= V\left[\frac{T^2}{n^2}\right] + \frac{\theta^4}{n^2} = \underbrace{\frac{1}{n^4}V[T^2] + \frac{\theta^4}{n^4}}$

∴ ① vs ②… $\underbrace{\frac{1}{n^2(n+1)^2}V[T^2]}_{小} < \underbrace{\frac{1}{n^4}V[T^2]}_{大} + \underbrace{\frac{\theta^4}{n^4}}_{多}$

∴ $\hat{\theta}_{UMVUE}$ 的 MSE 比較小

No.
Date

$\boxed{25}$

(a) $f(x|\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-x^2}{2\sigma^2}\right) I(x \in R)$

$= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma} \exp\left(\frac{-x^2}{2\sigma^2}\right) I(x \in R)$

$= \frac{1}{\sqrt{2\pi}} \cdot I(x \in R) \cdot \exp\left(\frac{-x^2}{2\sigma^2}\right) \cdot \exp(-\log \sigma)$

$= \underbrace{\frac{1}{\sqrt{2\pi}} I(x \in R)}_{h(x)} \cdot \exp\left(\underbrace{\frac{-1}{2\sigma^2}}_{\eta(\sigma)} \cdot \underbrace{x^2}_{T(x)} \underbrace{-\log \sigma}_{\psi(\sigma)}\right)$

$\therefore$ $h(x) = \frac{1}{\sqrt{2\pi}} I(x \in R)$

$\eta(\sigma) = \frac{-1}{2\sigma^2}$

$T(x) = x^2$

$\psi(\sigma) = \log \sigma$

(b) $f(x_1, x_2, \dots x_n | \sigma) = \prod_{j=1}^{n} f(x_j | \sigma)$

$= h(x_1) h(x_2) \dots h(x_n) \exp\left(\frac{-1}{2\sigma^2} \sum_{j=1}^{n} x_j^2 - n \log \sigma\right)$

No.
Date

$$\therefore \begin{cases} h(\vec{x}) = \left(\frac{1}{\sqrt{2\pi}}\right)^n I\left((x_1, \ldots, x_n) \in R^n\right) \\ \eta(\sigma) = -\frac{1}{2\sigma^2} \\ T_n(\vec{x}) = \sum_{j=1}^n x_j^2 \\ \psi_n(\sigma) = n \log \sigma \end{cases}$$

$\therefore$ 根据 Neyman-Fisher 的分解定理，

$T_n(\vec{x}) = \sum_{j=1}^n x_j^2$ 为 σ 之充份统计量，

(5) $\frac{x_j}{\sigma} \sim N(0,1)$ (iid)

根据 χ_n^2 的定义，$\sum_{j=1}^n \left(\frac{x_j}{\sigma}\right)^2 \sim \chi_n^2 = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$

$\therefore U \overset{def}{=} \sum_{j=1}^n \left(\frac{x_j}{\sigma}\right)^2 = \frac{T}{\sigma^2}$ $\qquad \frac{du}{dt} = \frac{1}{\sigma^2}$

$$f_U(u) = \frac{u^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right) \cdot 2^{\frac{n}{2}}} \exp\left(\frac{-u}{2}\right)$$

$$全概率 = 1 = \int_{u=0}^{u=\infty} \frac{u^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} \exp\left(\frac{-u}{2}\right) du \quad (u \to t)$$

$$= \int_{t=0}^{t=\infty} \frac{\left(\frac{t}{\sigma^2}\right)^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} \exp\left(\frac{-t}{2\sigma^2}\right) \cdot \frac{dt}{\sigma^2}$$

$$\therefore f_T(t) = \frac{t^{\frac{n}{2}-1}}{\sigma^n \Gamma\left(\frac{n}{2}\right) \cdot 2^{\frac{n}{2}}} \exp\left(\frac{-t}{2\sigma^2}\right) \quad (t > 0)$$

$$f_T(t) = \frac{t^{\frac{n}{2}-1}}{\sigma^n \Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} \exp\left(\frac{t}{2\sigma^2}\right) \quad (t>0)$$

$$= \underbrace{\frac{t^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right)\cdot 2^{\frac{n}{2}}}}_{h(t)} \cdot \exp\left(\underbrace{\frac{-1}{2\sigma^2}}_{\eta(\sigma)}\cdot \underbrace{t}_{s(t)} - \underbrace{n\log\sigma}_{\psi(\sigma)}\right)$$

高等統計推論 II

No ①

作業2、森元俊成　R05246013

1　$X \sim Bin(n,p)$

(1) $Pr(X=x) = {}_nC_x \, p^x (1-p)^{n-x} = L(p|x)$

$\log L(p|x) = (n-x)\log(1-p) + x\log p + \log {}_nC_x$

$\frac{\partial}{\partial p}\log L(p|x) = \frac{-1}{1-p}(n-x) + \frac{x}{p}$

$= \frac{x(1-p)-p(n-x)}{p(1-p)} = \frac{x-np}{p(1-p)} = 0$

$\Rightarrow p = \frac{x}{n}$

p		$\frac{x}{n}$	
L'	+	0	−
L	↗		↘

$\therefore p = \frac{x}{n} \Rightarrow L$: MAX

$\therefore \hat{p}_{MLE} = \frac{x}{n} \ \left(\frac{X}{n}\right)$

根據MLE的不變性，$p(1-p)$ 的MLE

為 $\hat{p}(1-\hat{p})$

(2) $E[\hat{p}(1-\hat{p})] = E\left[\frac{X}{n}\left(1-\frac{X}{n}\right)\right]$

$$E\left[\frac{X(n-X)}{n^2}\right] = \frac{1}{n^2} E[nX - X^2]$$

$$V[X] = E[X^2] - \underbrace{E[X]^2}_{(np)^2} = np(1-p)$$

$$\therefore\ E[X^2] = (np)^2 + np(1-p)$$

$$E[nX] = n^2 p$$

$$\frac{1}{n^2} E[nX - X^2] = \frac{1}{n^2}\left(n^2 p - (np)^2 - np(1-p)\right)$$

$$= \frac{1}{n^2}\left(n^2 p(1-p) - np(1-p)\right)$$

$$= \frac{1}{n^2}(n^2 - n)\, p(1-p)$$

$$= \left(1 - \frac{1}{n}\right) p(1-p)$$

$$\therefore\ E[\hat{p}(1-\hat{p})] = \left(1 - \frac{1}{n}\right) p(1-p)$$

∴ 並非 $p(1-p)$ 之不偏估計量

(3) $$E\left[\frac{1}{1-\frac{1}{n}}\, \hat{p}(1-\hat{p})\right] = p(1-p)$$

$$\therefore\ \frac{n}{n-1}\, \hat{p}(1-\hat{p})$$

$\boxed{2}$ $X \sim Bin(2,p)$

若X為完備, $E[g(X)]=0 \Rightarrow Pr(g(X)=0)=1$.

$E[X] = 2p$

$V[X] = 2p(1-p)$

$\therefore E[X^2] = 2p(1-p)+(2p)^2 = 2p^2+2p$

$E[aX^2+bX+c] = a(2p^2+2p)+2bp+c$

$$\begin{cases} E[aX^2+bX+c \mid p=\frac{1}{2}] = \frac{3}{2}a+b+c=0 & \cdots ① \\ E[aX^2+bX+c \mid p=\frac{1}{4}] = \frac{5a}{8}+\frac{b}{2}+c=0 & \cdots ② \end{cases}$$

$\therefore \frac{7a}{8}+\frac{b}{2}=0 \quad (\because ①-②)$

$\therefore b=\frac{-7}{4}a$

$\therefore a=1 \Rightarrow b=\frac{-7}{4} \quad c=\frac{1}{4}$

$\therefore E[X^2-\frac{7X}{4}+\frac{1}{4} \mid p]=0 \quad (\text{for all } p=\frac{1}{2} \text{ or } \frac{1}{4})$

$g(X)=X^2-\frac{7X}{4}+\frac{1}{4} \quad \begin{cases} Pr(g(X)=0) \neq 1 \\ E[g(X)]=0 \end{cases}$

$\therefore$ X 並非完備

No.

Date

$\boxed{3}$ $U_1 \stackrel{def}{:=} \frac{X_1}{\theta}$, $U_2 \stackrel{def}{=} \frac{X_2}{\theta}$ $U_3 \stackrel{def}{=} \frac{X_3}{\theta}$

$U_1, U_2, U_3 \sim U(0,1)$. 顯然 $U_{(1)} = \frac{X_{(1)}}{\theta}$.

(a) $Pr(X_{(1)} \leq x) = 1 - Pr(X_{(1)} > x) = 1 - Pr(X_1 \sim X_3 > x)$

$= 1 - \left(\frac{\theta - x}{\theta}\right)^3 = 1 - \left(1 - \frac{x}{\theta}\right)^3 \quad (0 \leq x \leq \theta)$

$\frac{d}{dx} Pr(X_{(1)} \leq x) = -3\left(1 - \frac{x}{\theta}\right)^2 \cdot \left(\frac{-1}{\theta}\right) = \frac{3}{\theta}\left(1 - \frac{x}{\theta}\right)^2 \quad (0 \leq x \leq \theta)$

$0 \quad$ (otherwise)

$\therefore f_{X_{(1)}}(x) = \frac{3}{\theta}\left(1 - \frac{x}{\theta}\right)^2 \cdot I_{(0,\theta)}(x)$

$U_{(1)} = \frac{X_{(1)}}{\theta} \quad (U_1 \sim U_3 \sim U(0,1))$

$\therefore f_{X_{(1)}}(x \mid \theta = 1) = 3(1-x)^2 \underset{I_{(0,1)}(x)}{=} \frac{1}{Be(1,3)} \cdot x^{1-1}(1-x)^{3-1} \cdot I_{(0,1)}(x)$

由此可知, $\frac{X_{(1)}}{\theta} \sim Be(1,3)$

(b) $E[X_{(1)}] = \theta \cdot E\left[\frac{X_{(1)}}{\theta}\right] = \theta \cdot E[U_{(1)}]$

$= \theta \cdot \int_{x=0}^{x=1} 3x(1-x)^2 dx = 3\theta \cdot Be(2,3)$

$= 3\theta \cdot \frac{\Gamma(2)\Gamma(3)}{\Gamma(5)} = 3\theta \cdot \frac{2!}{4!} = \frac{\theta}{4}$

⑤

$E[X_{(1)}] = \frac{\theta}{4}$ ∴ $E[4X_{(1)}] = \theta$

∴ $4X_{(1)}$

(c) 將情況一般化：

$X_1, X_2, \ldots, X_n \sim Uni(0, \theta)$

證明 $\frac{X_{(n)}}{X_{(1)}}$ vs $X_{(n)}$ 相互獨立

① $X_{(n)}$ 為 θ 之 完備充份統計量

原因：$\frac{d}{dx} Pr(X_{(n)} \leq x) = \frac{d}{dx} Pr(X_1 \sim X_n \leq x)$

$= \frac{d}{dx}\left(\frac{x}{\theta}\right)^n = \frac{nx^{n-1}}{\theta^n} \quad (0 \leq x \leq \theta)$

$0 \quad (\text{elsewhere})$

$f_{X_{(n)}}(x) = \frac{nx^{n-1}}{\theta^n} \cdot I_{(0,\theta)}(x) \quad (\theta > 0)$

$E[g(X)] = 0 \Rightarrow \int_0^\theta g(x) \frac{nx^{n-1}}{\theta^n} dx = 0$

$\Rightarrow \int_0^\theta g(x) \cdot x^{n-1} dx = 0 \Rightarrow \frac{d}{d\theta} \int_0^\theta g(x) x^{n-1} dx = 0$

$\Rightarrow g(\theta) \cdot \theta^{n-1} = 0 \Rightarrow g(\theta) = 0 \quad (\text{for all } \theta \in \{\theta \mid \theta > 0\})$

$(\theta > 0)$

∴ $Pr(g(X) = 0) = 1$

$\therefore X_{(n)}$ 為 θ 之完備充份統計量

② $\frac{X_{(n)}}{X_{(1)}}$ 為 θ 之輔助統計量

$$\because \frac{X_{(n)}}{X_{(1)}} = \frac{\left(\frac{X_{(n)}}{\theta}\right)}{\left(\frac{X_{(1)}}{\theta}\right)} \quad \left(\frac{X_{(n)}}{\theta}, \frac{X_{(1)}}{\theta} \text{ 的分佈與 } \theta \text{ 無關}\right)$$

$\downarrow$

(跟 θ 無關)

③ BASU 定理…… $\underbrace{\frac{X_{(n)}}{X_{(1)}}}_{\text{輔助}}$ vs $\underbrace{X_{(n)}}_{\text{完備}}$ 為獨立

證) $S = \frac{X_{(n)}}{X_{(1)}}$, $T = X_{(n)}$

$S|T$ 的分佈與 θ 無關 ($\because T$ = 充份統計量)

$Pr(S\in A|T)$ …… 與 θ 無關, …… T 的函數

$$E[Pr(S\in A|T)] = \int_T \underbrace{Pr(S\in A|T)}_{\shortparallel} \cdot f_T(t)\,dt$$

$$\int_{S\in A} f_{S|T}(s|t)\,ds$$

$$= \int_T \int_{S\in A} \frac{f_{S,t}(s,t)}{f_T(t)}\,ds \cdot f_T(t)\,dt = \int_T \int_{S\in A} f_{S,t}(s,t)\,ds\,dt$$

$= Pr(S\in A)$ …… 常數

No.
Date

$E[\Pr(S\in A|T)] = \Pr(S\in A)$

$E[\underbrace{\Pr(S\in A|T) - \Pr(S\in A)}_{=\,T\text{(完備充分統計量)的函數}}] = 0$

根據完備充分統計量的定義,

$\Pr(\Pr(S\in A|T) - \Pr(S\in A) = 0) = 1$

$\therefore \Pr(S\in A|T) = \Pr(S\in A)$ (a.s) (for all A)

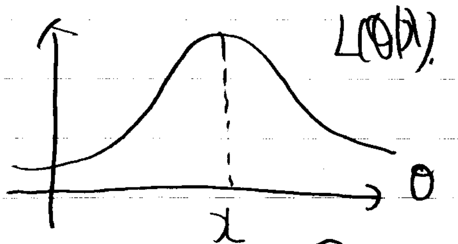
∴ 由此可知, S, T 為獨立。

$\boxed{4}$ $X \sim N(\theta, 1)$ $(\theta > 0)$

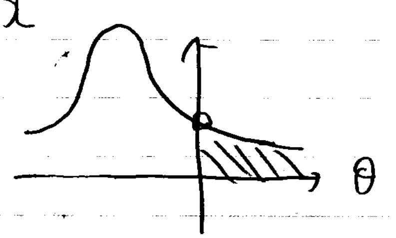
(a) $f(x|\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2\pi}(x-\theta)^2\right)$

$\because f(x|\theta) = f(\theta|x)$, $L(\theta|x) \overset{def}{=} f(x|\theta)$

∴



① if $x < 0 \Rightarrow$



Chryo culture

No.
Date

若允許 $\theta=0 \Rightarrow$ $\theta=0$ 使得 $L(\theta|x)$ 為最大

$\hat{\theta}_{MLE}=0$

(不允許 $\Rightarrow$ MLE 不存在)

② $x \geq 0$、 $\theta=x$ 使得 $L(\theta|x)$ 為最大

$\hat{\theta}_{MLE}=x$

∴ 允許 $\theta=0 \Rightarrow \hat{\theta}_{MLE}=\max\{x,0\}$

(不允許 $\theta=0 \Rightarrow \hat{\theta}_{MLE}=\begin{cases} x & (x\geq 0) \\ \text{不存在} \end{cases}$)

(b) 假設 $\hat{\theta}_{MLE}=\max\{x,0\}$ (換言之允許 $\theta=0$)

$F_{\hat{\theta}_{MLE}}(t)=\Pr(\hat{\theta}_{MLE}\leq t)=\Pr(X\leq t \cap 0\leq t)$
(cdf)

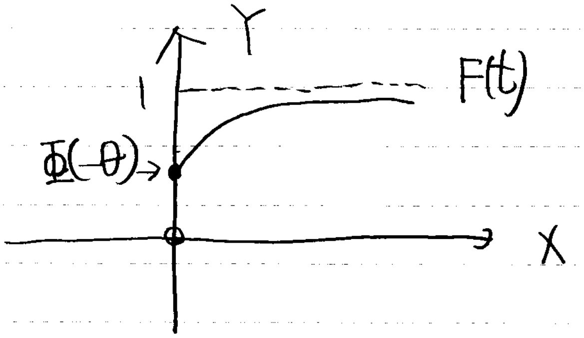
$=\Pr(X-\theta\leq t-\theta \cap 0\leq t)$

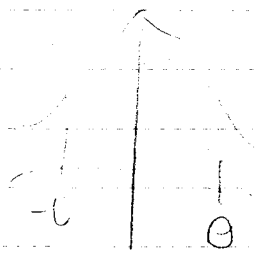
$=\Pr(X-\theta\leq t-\theta)\cdot I_{[0,\infty)}(t)$

$=\Phi(t-\theta)\cdot I_{[0,\infty)}(t)$

⑨

$$F_{\hat{\theta}_{MLE}}(t) = \Phi(t-\theta)\cdot I_{[0,\infty)}(t)$$





∴ $\hat{\theta}_{MLE}$ 包含一個 atom. ($\hat{\theta}_{MLE}=0$)

① $Pr(\hat{\theta}_{MLE}=0) = \Phi(-\theta) = 1-\Phi(\theta)$ （離散的部分）

② $\hat{\theta}_{MLE}>0 \Rightarrow \frac{d}{dt}F_{\hat{\theta}_{MLE}}(t) = \phi(t-\theta)$ （連續的部分）

$$\therefore E[\hat{\theta}_{MLE}] = \overset{①}{0\cdot Pr(\hat{\theta}_{MLE}=0)} + \overset{②}{\int_{t>0} t\,\phi(t-\theta)\,dt}$$

$$= \int_{t>0} t\,\phi(t-\theta)\,dt \qquad (u \overset{put}{=} t-\theta)$$

$$= \int_{-\theta}^{\infty} (u+\theta)\,\phi(u)\,du$$

$$= \underbrace{\int_{-\theta}^{\infty} u\,\phi(u)\,du}_{} + \theta\underbrace{\int_{-\theta}^{\infty} \phi(u)\,du}_{\parallel}$$

$$\left[\Phi(u)\right]_{-\theta}^{\infty} = 1-\Phi(-\theta) = \Phi(\theta)$$

$$\left[-\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-u^2}{2}\right)\right]_{-\theta}^{\infty} = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{-\theta^2}{2}\right)$$

$$\int_t^{\infty} \lambda \exp(-\lambda x)\, dx = \left[\exp(-\lambda x)\right]_t^{\infty} = \exp$$

$$\therefore E[\hat{\theta}_{MLE}] = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\theta^2}{2}\right) + \theta\,\Phi(\theta)$$

$$\therefore E[\hat{\theta}_{MLE}] \neq \theta$$

$\boxed{5}$ $X \sim \exp(\lambda)$ $Y \sim \exp(\lambda)$ $(\alpha \to \lambda)$
(independent)

① $$Pr(X^3 \leq t) = Pr(X \leq t^{\frac{1}{3}}) = 1 - Pr(X > t^{\frac{1}{3}})$$

$$= 1 - \exp(-\lambda t^{\frac{1}{3}}) \quad (t > 0)$$

$$\frac{d}{dt} Pr(X^3 \leq t) = \frac{\lambda}{3} t^{-\frac{2}{3}} \exp(-\lambda t^{\frac{1}{3}}) \quad (t > 0)$$

$$\therefore f_{X^3}(t) = \frac{\lambda}{3} t^{-\frac{2}{3}} \exp(-\lambda t^{\frac{1}{3}}) \quad (t > 0)$$

② 將情況一般化...

$|X - Y|$ = 樣本範圍 (Sample Range)

$X_1 \dots X_n \overset{\text{iid}}{\sim} e(\lambda)$

考慮 $X_{(n)} - X_{(1)}$ 之分佈

$$\begin{cases} R = X_{(n)} - X_{(1)} \\ S = \dfrac{X_{(n)} + X_{(1)}}{2} \end{cases}$$

(11)

$Pr(A \cap B) = Pr(B$

$$Pr(X_{(1)} \leqq x, X_{(n)} \leqq y) = Pr(X_{(n)} \leqq y) - Pr(x < X_{(1)}, X_{(n)} \leqq y)$$

$$= \underbrace{Pr(X_1 \wedge X_n \leqq y)} - Pr(x < X_1 \wedge X_n \leqq y)$$

$$(1 - \exp(-\lambda y))^n - (\exp(-\lambda x) - \exp(-\lambda y))^n$$

$$\frac{\partial^2}{\partial x \partial y} Pr(X_{(1)} \leqq x, X_{(n)} \leqq y) =$$

$$\frac{\partial}{\partial y} -n(\exp(-\lambda x) - \exp(-\lambda y))^{n-1} \cdot (-\lambda)\exp(-\lambda x) =$$

$$\frac{\partial}{\partial y} \{ n\lambda \exp(-\lambda x)(\exp(-\lambda x) - \exp(-\lambda y))^{n-1}$$

$$= n(n-1)\lambda^2 \exp(-\lambda x)\exp(-\lambda y)(\exp(-\lambda x) - \exp(-\lambda y))^{n-2} \quad (x \geqq 0, y \geqq 0)$$

$$\therefore f_{X_{(1)}, X_{(n)}}(x, y) = n(n-1)\lambda^2 \exp(-\lambda(x+y))\{\exp(-\lambda x) - \exp(-\lambda y)\}^{n-2}$$

$$(x \geqq 0, y \geqq 0)$$

$$X_{(n)} = S + \frac{r}{2} \quad (= y)$$

$$X_{(1)} = S - \frac{r}{2} \quad (= x)$$

$$\therefore 0 \leqq S - \frac{r}{2} \leqq S + \frac{r}{2} < \infty$$

$$\therefore 0 \leqq \frac{r}{2} \leqq S < \infty$$

No.
Date

$$\frac{\partial}{\partial s}\begin{pmatrix} X_{(n)} \\ X_{(1)} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \frac{\partial}{\partial r}\begin{pmatrix} X_{(n)} \\ X_{(1)} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$$

$\therefore |J| = 1 \qquad \therefore dx_{(n)}\,dx_{(1)} = dr\,ds$

$\therefore$ 全確率 $= 1 = \iint_{0 \leq X_{(1)} \leq X_{(n)} < \infty} f_{X_{(1)},X_{(n)}}(x,y)\,dx\,dy$

$$= \iint_{0 \leq \frac{r}{2} \leq s < \infty} n(n-1)\lambda^2 \exp(-2\lambda s) \cdot \left\{ \exp\left(-\lambda s + \frac{\lambda r}{2}\right) - \exp\left(-\lambda s - \frac{\lambda r}{2}\right) \right\}^{n-2} dr\,ds$$

$$= \iint_{0 \leq \frac{r}{2} \leq s < \infty} n(n-1)\lambda^2 \exp(-2\lambda s) \cdot \left\{ \exp(-\lambda s) \cdot \left(\exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right) \right) \right\}^{n-2} dr\,ds$$

$$= \iint_{(\approx)} n(n-1)\lambda^2 \exp(-2\lambda s) \exp(-\lambda(n-2)s) \left\{ \exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right) \right\}^{n-2} dr\,ds$$

$$= \int_{0 \leq \frac{r}{2} \leq s < \infty} n(n-1)\lambda^2 \exp(-\lambda n s) \left\{ \exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right) \right\}^{n-2} dr\,ds$$

求 R 的邊際分布 $\Rightarrow \int_{s \geq \frac{r}{2}} \exp(-\lambda n s)\,ds$

$$= \left[\frac{-1}{\lambda n} \exp(-\lambda n s) \right]_{\frac{r}{2}}^{\infty} = \frac{1}{\lambda n} \exp\left(\frac{-n\lambda r}{2}\right)$$

$$\therefore f_R(r) = \cancel{n}(n-1)\lambda^2 \cdot \frac{1}{\cancel{\lambda n}} \exp\left(\frac{-n\lambda r}{2}\right) \left\{ \exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right) \right\}^{n-2} \quad (r \geq 0)$$

$$= (n-1)\lambda \exp\left(\frac{-n\lambda r}{2}\right) \left\{ \exp\left(\frac{\lambda r}{2}\right) - \exp\left(\frac{-\lambda r}{2}\right) \right\}^{n-2} \quad (r \geq 0)$$

$\therefore n = 2$ 時, $f_R(r) = \lambda \exp(-\lambda r) \quad \therefore R \sim \exp(\lambda)$ $\left(\text{mean} = \frac{1}{\lambda}\right)$

Chuyu culture

⑬

(2) $\Pr(\min\{X, Y^3\} \le t) = 1 - \Pr(\min\{X, Y^3\} > t)$

$= 1 - \Pr(X, Y^3 > t) = 1 - \Pr(X > t)\Pr(Y^3 > t)$

$= 1 - \Pr(X > t)\Pr(Y > t^{\frac{1}{3}}) \quad (t > 0)$

$= 1 - \exp(-\lambda t) \cdot \exp(-\lambda t^{\frac{1}{3}})$

$= 1 - \exp(-\lambda(t + t^{\frac{1}{3}}))$

$\frac{d}{dt}\Pr(\min\{X, Y^3\} \le t) = \frac{d}{dt}(1 - \exp(-\lambda(t + t^{\frac{1}{3}})))$

$= -\frac{d}{dt}(-\lambda(t + t^{\frac{1}{3}})) \cdot \exp(-\lambda(t + t^{\frac{1}{3}}))$

$= \lambda(1 + \frac{1}{3}t^{-\frac{2}{3}})\exp(-\lambda(t + t^{\frac{1}{3}})) \quad (t > 0)$

$$\therefore f_{\min\{X, Y^3\}}(t) = \begin{cases} \lambda(1 + \frac{1}{3}t^{-\frac{2}{3}})\exp(-\lambda(t + t^{\frac{1}{3}})) & (t > 0) \\ 0 & (\text{else}) \end{cases}$$

No

Date

6 證明 $E[rU]=\frac{r}{\alpha}$ $V[rU]=\frac{r}{\alpha^2}$ 即可

先考慮 $X(1) \cdots X(r)$ 的聯合分佈

$$f(X(1) \cdots X(n)) = n!\, \alpha^n \exp(-\alpha X(1) - \alpha X(2) \cdots -\alpha X(n))$$

$$(0 \leqq X(1) \leqq X(2) \cdots \leqq X(n))$$

①

$$\int_{X(n) \geqq X(n-1)} f(X(1) \cdots X(n))\, dX(n)$$

$$= n!\, \alpha^{n-1} \exp(-\alpha X(1)) \cdots \exp(-\alpha X(n-1)) \cdot \Big[-\exp(-\alpha X(n))\Big]_{X(n-1)}^{\infty}$$

$$= n!\, \alpha^{n-1} \exp(-\alpha X(1)) \cdots \exp(-\alpha X(n-2)) \cdot \exp(-2\alpha X(n-1))$$

(↑ $X(1) \sim X(n-1)$ 之聯合分布)

② $$\int_{X(n-1) \geqq X(n-2)} f(X(1) \cdots X(n-1))\, dX(n-1)$$

$$= n!\, \alpha^{n-2} \exp(-\alpha X(1)) \cdots \exp(-\alpha X(n-2)) \cdot \left[\frac{-1}{2} \exp(-2\alpha X(n-1))\right]_{X(n-2)}^{\infty}$$

$$= \frac{n!}{2!} \alpha^{n-2} \exp(-\alpha X(1)) \cdots \exp(-\alpha X(n-3)) \cdot \exp(-3\alpha X(n-2))$$

(↑ $X(1) \sim X(n-2)$ 之聯合分布)

⫶

ⓙ $$\frac{n!}{j!} \alpha^{n-j} \exp(-\alpha X(1)) \cdots \exp(-\alpha X(n-j-1)) \exp(-(j+1)\alpha X(n-j))$$

(↑ $X(1) \sim X(n-j)$ 之聯合分布)

Chuyu culture

∴ $\int^{\text{卅}} n-r$ 即可得 $X_{(1)} \sim X_{(r)}$ 之聯合分佈

$$\frac{n!}{(n-r)!}\alpha^{r}\exp(-\alpha X_{(1)})\cdots\exp(-\alpha X_{(r-1)})\exp\left(-(n-r+1)\alpha X_{(r)}\right)$$

$$\left(0\le X_{(1)}\le\cdots\le X_{(r)}\right)$$

接下來考慮以下變數轉換

$$\begin{cases} Y_1 = nX_{(1)} \\ Y_2 = (n-1)(X_{(2)}-X_{(1)}) \\ \vdots \\ Y_r = (n-r+1)(X_{(r)}-X_{(r-1)}) \end{cases}$$

$\frac{Y_{(3)}}{n-2} = X_{(3)} - X_{(2)}$

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_r \end{pmatrix} = \begin{pmatrix} n & & & & 0 \\ -(n-1) & (n-1) & & & \\ & -(n-2) & (n-2) & & \\ & & & \ddots & \\ 0 & & & -(n-r+1) & (n-r+1) \end{pmatrix} \begin{pmatrix} X_{(1)} \\ \vdots \\ X_{(r)} \end{pmatrix}$$

$$\begin{aligned} dy_1\cdots dy_r &= |J|\, dx_1\cdots dx_r \\ &= n(n-1)\cdots(n-r+1)\, dx_1\cdots dx_r \\ &= \frac{n!}{(n-r)!}\, dx_1\cdots dx_r. \end{aligned}$$

No.
Date

$$X_{(1)} = \frac{Y_1}{n},\quad X_{(2)} = \frac{Y_1}{n} + \frac{Y_2}{n-1}\quad X_{(3)} = \frac{Y_1}{n} + \frac{Y_2}{n-1} + \frac{Y_3}{n-2} \cdots$$

$$0 \leqq X_{(1)} \leqq X_{(2)} \leqq \cdots$$

$$\Leftrightarrow\ Y_1 \geqq 0,\ Y_2 \geqq 0 \cdots\ Y_r \geqq 0$$

全概率

$$1 = \iint\cdots\int_{0 \leqq X_{(1)} \leqq \cdots \leqq X_{(r)}} \frac{n!}{(n-r)!} \alpha^r \exp(-\alpha X_{(1)}) \cdots \exp(-(n-r+1) X_{(r)})\, dx_{(1)} \cdots dx_{(r)}$$

(另外 $Y_1 + Y_2 + \cdots Y_r = X_{(1)} + X_{(2)} + \cdots X_{(r-1)} + (n-r+1) X_{(r)}$)

全概率

$$\therefore 1 = \int\cdots\int_{Y_1 \geqq 0 \sim Y_r \geqq 0} \frac{n!}{(n-r)!} \alpha^r \exp(-\alpha(Y_1 + Y_2 + \cdots + Y_r)) \cdot \frac{(n-r)!}{n!}\, dy_1 \cdots dy_n$$

$$= \int\cdots\int_{Y_1 \geqq 0 \sim Y_r \geqq 0} \alpha \exp(-\alpha y_1) \cdot \alpha \exp(-\alpha y_2) \cdots \alpha \exp(-\alpha y_r)\, dy_1 \cdots dy_n$$

由此可知 $Y_1 \sim Y_n \overset{\text{i.i.d.}}{\sim} \exp(\alpha)$ (mean : $\frac{1}{\alpha}$)

$$\therefore Y_1 + Y_2 + \cdots + Y_r = X_{(1)} + X_{(2)} + \cdots X_{(r-1)} + (n-r+1) X_{(r)}$$

$$= rU \sim \Gamma(r, \alpha)\ \text{(Gamma分布)}$$

$\therefore E[rU] = \frac{r}{\alpha}$, $V[rU] = \frac{r}{\alpha^2}$ ($\because$ Gamma分佈) $\therefore$ 證明完了

⊛ 7. 和 8. 的中間還有一題，所以 $\boxed{7}$ 分成 $\boxed{7\text{-}1}$ 和 $\boxed{7\text{-}2}$

$\boxed{7\text{-}1}$ $X_1 \sim X_n$; iid $f(x) = \dfrac{\theta}{(1+x)^{\theta+1}}$ $(x>0)$ $(\theta>0)$

(a) $f(x_1 \cdots x_n \mid \theta) = \dfrac{\theta}{(1+x_1)^{\theta+1}} \cdots \dfrac{\theta}{(1+x_n)^{\theta+1}}$ $(x_1 \sim x_n > 0)$

$$L(\theta \mid \vec{x}) = \log f(x_1 \cdots x_n \mid \theta) = \log \frac{\theta^n}{(1+x_1)^{\theta+1} \cdots (1+x_n)^{\theta+1}}$$

$$= n \log\theta - (\theta+1)\sum_{j=1}^{n} \log(1+x_j)$$

$$\frac{\partial}{\partial\theta} L(\theta \mid \vec{x}) = \frac{n}{\theta} - \sum_{j=1}^{n} \log(1+x_j) = 0$$

$$\therefore \frac{\theta}{n} = \frac{1}{\sum_{j=1}^{n} \log(1+x_j)}$$

$$\therefore \theta = \frac{n}{\sum_{j=1}^{n} \log(1+x_j)}$$

θ	0		$\sum^{n} \log(1+x_j)$	
$L'(\theta)$		$+$	0	$-$
$L(\theta)$		↗	max	↘

$$\therefore \hat{\theta}_{MLE} = \frac{n}{\sum_{j=1}^{n} \log(1+x_j)}$$

No.
Date

(b) ① 求 $\lg(1+X_j)$ 的分布
② $\frac{\sum \lg(1+X_j)}{n}$ … 中央極限定理
③ $g(x)$一式 … δ-method.

$$Y_j \overset{def}{=} \lg(1+X_j) \qquad \begin{matrix} x_j = 0 \to \infty \\ y_j = 0 \to \infty \end{matrix}$$

$$\frac{dy_j}{dx_j} = \frac{1}{1+x_j} = e^{-y_j}$$

$$全概率 = 1 = \int_{x_j=0}^{x_j=\infty} \frac{\theta}{(1+x_j)^{\theta+1}} dx_j$$

$$= \int_{y_j=0}^{y_j=\infty} \theta e^{-(\theta+1)y_j} \cdot e^{y_j} dy_j$$

$$= \int_0^\infty \theta e^{-\theta y_j} dy_j$$

由此可知 $Y_j \sim exp(\theta)$ (mean: $\frac{1}{\theta}$, var: $\frac{1}{\theta^2}$)

據此 $E[Y_j^2] < \infty$ ∴ 根據中央極限定理

$$\frac{Y_1 + \cdots + Y_n}{n} \xrightarrow{d} N\left(\frac{1}{\theta}, \frac{1}{n\theta^2}\right)$$

(19)

$$\therefore \sqrt{n}\left(\bar{Y}-\frac{1}{\theta}\right) \xrightarrow{d} N\left(0, \frac{1}{\theta^2}\right)$$

$$g(x) = \frac{1}{x}$$

$$g(\bar{Y}) \approx g'\left(\frac{1}{\theta}\right)\left(\bar{Y}-\frac{1}{\theta}\right) + g\left(\frac{1}{\theta}\right)$$

$$\therefore g(\bar{Y}) - g\left(\frac{1}{\theta}\right) \approx g'\left(\frac{1}{\theta}\right)\left(\bar{Y}-\frac{1}{\theta}\right)$$

$$\sqrt{n}\left(g(\bar{Y}) - g\left(\frac{1}{\theta}\right)\right) \approx g'\left(\frac{1}{\theta}\right)\underbrace{\sqrt{n}\left(\bar{Y}-\frac{1}{\theta}\right)}_{\xrightarrow{d} N\left(0, \frac{1}{\theta^2}\right)}$$

$$\therefore \xrightarrow{d} N\left(0, \left(g'\left(\frac{1}{\theta}\right)\right)^2 \cdot \frac{1}{\theta^2}\right)$$

$$g'(x) = -\frac{1}{x^2} \qquad g'\left(\frac{1}{\theta}\right) = -\theta^2$$

$$= N\left(0, \theta^4 \cdot \frac{1}{\theta^2}\right)$$

$$\| $$

$$N(0, \theta^2)$$

$$\therefore \sqrt{n}\left(g(\bar{Y}) - g\left(\frac{1}{\theta}\right)\right) \xrightarrow{d} N(0, \theta^2)$$

$$\frac{1}{\bar{Y}} = \frac{n}{\sum_{i=1}^{n} \log(1+X_i)} = \hat{\theta}_{MLE}$$

No.
Date

$\therefore \sqrt{n}(\hat{\theta}_{MLE}-\theta) \xrightarrow{d} N(0,\theta^2)$

(c) 同樣利用 δ-method

$$\sqrt{n}(g(\hat{\theta}_{MLE})-g(\theta)) \xrightarrow{d} N(0, \{g'(\theta)\}^2 \cdot \theta^2)$$
$$= N(0,1)$$

$$\therefore g'(\theta) = \frac{\pm 1}{\theta} + c$$

$$\therefore g(\theta) = \pm \log\theta + c$$

(c: 常數)

7和8之間有一題

7-2

(a) $\hat{p}_{IS} = \frac{1}{m}\sum_{i=1}^{m} I(x_i \geq 5) \cdot \frac{\phi(x_i)}{\phi(x_i - 5)}$

$$E[\hat{p}_{IS}] = E\left[\frac{1}{m}\sum_{i=1}^{m} I(x_i \geq 5) \cdot \frac{\phi(x_i)}{\phi(x_i-5)}\right]$$

$$= \frac{1}{m}\sum_{i=1}^{m} \underbrace{E\left[I(x_i \geq 5) \cdot \frac{\phi(x_i)}{\phi(x_i-5)}\right]}$$

$$\int_{x_i \geq 5} \frac{\phi(x_i)}{\phi(x_i-5)} \cdot \underbrace{\phi(x_i-5)}_{x_i\text{之pdf}}\, dx_i$$

$$= \int_{x_i \geq 5} \phi(x_i)\, dx_i$$

$$= \left[\Phi(x_i)\right]_5^{\infty} = 1 - \Phi(5)$$

$$= \frac{1}{m}\sum_{i=1}^{m} (1 - \Phi(5)) = \underline{1 - \Phi(5)}$$

$\therefore \hat{p}_{IS}$ 為 $1 - \Phi(5)$ 之不偏估計量

(b) $$V[\hat{p}_{IS}] = V\left[\frac{1}{m}\sum_{i=1}^{m} I(x_i \geq 5) \frac{\phi(x_i)}{\phi(x_i-5)}\right]$$

$$= \frac{1}{m^2}\sum_{i=1}^{m} V\left[I(x_i \geq 5) \frac{\phi(x_i)}{\phi(x_i-5)}\right]$$

$$= \frac{1}{m} V\left[I(x_i \geq 5) \cdot \frac{\phi(x_i)}{\phi(x_i-5)}\right]$$ (∵ 獨立)

No.

Date

$$E\left[\left(I(X_1 \geq 5)\cdot\frac{\phi(X_1)}{\phi(X_1-5)}\right)^2\right]$$

$$= \int_{x_1 \in R} I(X_1 \geq 5)\cdot\frac{\phi(x_1)^2}{\phi(x_1-5)^2}\cdot\phi(x_1-5)\,dx_1$$

$$= \int_{x_1 \geq 5} \frac{\phi(x_1)^2}{\phi(x_1-5)}\,dx_1$$

$$\because \frac{\phi(x_1)^2}{\phi(x_1-5)} = \frac{1}{2\pi}\exp(-x_1^2)\cdot\sqrt{2\pi}\exp\left(\frac{1}{2}(x_1-5)^2\right)$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(-x_1^2+\frac{1}{2}(x_1^2-10x_1+25)\right)$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x_1^2-5x_1+\frac{25}{2}\right)$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x_1+5)^2+\frac{50}{2}\right)$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x_1+5)^2\right)\cdot\exp(25)$$

$$\therefore = \int_{x_1 \geq 5} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(x_1+5)^2\right)\exp(25)\,dx_1$$

$$z = x_1+5 \qquad dx_1 = dz$$

$$= \int_{z \geq 10} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}z^2\right)\cdot\exp(25)\,dz$$

$$= \exp(25)\left(1-\Phi(10)\right)$$

Chryu culture

22

$$\therefore E\left[\left\{ I(X_1 \geq 25) \cdot \frac{\phi(X_1)}{\phi(X_1-5)} \right\}^2\right] = e^{?(25)}\left(1-\Phi(10)\right)$$

$$\therefore V\left[I(X_1 \geq 25) \frac{\phi(X_1)}{\phi(X_1-5)} \right] = e^{?(25)}\left(1-\Phi(10)\right) - \left(1-\Phi(5)\right)^2$$

$$\therefore V[\hat{p}_{25}] = \frac{1}{m}\left\{ e^{?(25)}\left(1-\Phi(10)\right) - \left(1-\Phi(5)\right)^2 \right\}$$

$\boxed{8}$ $g(x) = \dfrac{f(x)}{1-F(x_0)} I_{[x_0,\infty)}(x)$

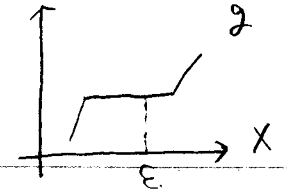
① $g(x) \geq 0$ $\left(\because f(x) > 0,\ 1-F(x_0) > 0\right)$

② $\displaystyle\int_{x \in R} g(x)\,dx = \int_{x \geq x_0} \frac{f(x)}{1-F(x_0)}\,dx$

$$= \frac{1}{1-F(x_0)}\left[F(x)\right]_{x_0}^{\infty} = \frac{1-F(x_0)}{1-F(x_0)} = 1$$

∴ 由①,②, $g(x)$ 爲 機率密度函數.

cdf: $G(x) = \displaystyle\int_{-\infty}^{x} g(t)\,dt = \begin{cases} \dfrac{F(x)-F(x_0)}{1-F(x_0)} & (x \geq x_0) \\ 0 & (x < x_0) \end{cases}$



$$\{\omega \mid X(\omega) \geq \varepsilon\} \subseteq \{\omega \mid g(X) \geq g(\varepsilon)\}$$

(24)

$\boxed{9}$ 由於 g 為遞增函數, $Pr(X \geq \varepsilon) \leq Pr(g(X) \geq g(\varepsilon))$ …… ※1

$$E[g(X)] = \int_{\omega \in \Omega} g(X) \cdot P(dX(\omega))$$

$$\geq \int_{\{\omega \in \Omega \mid g(X) \geq g(\varepsilon)\}} g(X)\, P(dX(\omega))$$

$$\begin{cases} \Omega \supseteq \{\omega \in \Omega \mid g(X) \geq g(\varepsilon)\} \\ g \cdots \text{non-negative} \end{cases}$$

$$\geq \int_{\{\omega \in \Omega \mid g(X) \geq g(\varepsilon)\}} g(\varepsilon)\, P(dX(\omega))$$

$$= g(\varepsilon) \int_{\{\omega \in \Omega \mid g(X) \geq g(\varepsilon)\}} P(dX(\omega))$$

$$= g(\varepsilon) \cdot P(\{\omega \in \Omega \mid g(X) \geq g(\varepsilon)\})$$

$$\geq g(\varepsilon) \cdot P(\{\omega \in \Omega \mid X \geq \varepsilon\}) \quad (\because ※1)$$

$$= g(\varepsilon)\, P(X \geq \varepsilon)$$

$$\therefore P(X \geq \varepsilon) \leq \frac{E[g(X)]}{g(\varepsilon)}$$

∴ 證明完成

- $\boxed{10}$ 跟 $\boxed{4}$ 一樣，請參閱 $\boxed{4}$。
- $\boxed{11}$ $T|N \sim N(N\mu, N\sigma^2)$ $\quad N \sim Po(\lambda)$

① $\sqrt{n}\left(\frac{T_1+\cdots T_n}{n} - \lambda\mu\right)$ 之極限分佈…

求 T_j 的期望值與變異數（$T_1, T_2 \cdots T_n$ iid 某種分佈）

- $E[T|N] = N\cdot\mu$

$E[E[T|N]] = E[N\mu] = \mu E[N] = \mu\lambda$

- $E[T^2|N] = N\sigma^2 + N^2\mu^2$

$E[E[T^2|N]] = E[N\sigma^2 + N^2\mu^2] = \lambda\sigma^2 + (\lambda^2+\lambda)\mu^2$

(註) $E[N] = \lambda$ $\quad E[N^2] = \lambda^2+\lambda$

故 $E[T] = \mu\lambda$ $\quad E[T^2] = \lambda\sigma^2 + (\lambda^2+\lambda)\mu^2$

$\Rightarrow V[T] = \lambda\sigma^2 + (\lambda^2+\lambda)\mu^2 - \mu^2\lambda^2$

$= \lambda\sigma^2 + \lambda\mu^2 = \lambda(\sigma^2+\mu^2)$ (for all T_j, $j=1 \sim n$)

$\therefore E[T_j] < \infty$, $V[T_j] < \infty$ （→ 中央極限定理）

根據中央極限定理，($T_1 \sim T_n$ 之 iid)

$$\sqrt{n}\left(\frac{T_1+T_2+\cdots+T_n}{n}-\lambda\mu\right) \xrightarrow{d} N(0, \lambda(\sigma^2+\mu^2))$$

② $\frac{\sqrt{n}}{N}(\bar{T}-\lambda\mu)$ 之極限分佈

利用 Slutsky 定理，$\begin{cases} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{p} c \ (\text{常數}) \end{cases}$

$$\Rightarrow X_n Y_n \xrightarrow{d} cX$$

我們已知 $\sqrt{n}(\bar{T}-\lambda\mu) \xrightarrow{d} N(0, \lambda(\sigma^2+\mu^2))$

觀察 $\frac{1}{N} = \frac{n}{N_1+N_2+\cdots N_n}$ 是否機率收斂到常數.

根據 Khinchine 弱大數法則. $\frac{N_1+\cdots N_n}{n} \xrightarrow{p} \lambda$ ($\because E[N]<\infty$)

由於 $g(x)=\frac{1}{x}$ ($\lambda>0$) 為 R 上的連續函數.

因此 $g\left(\frac{N_1+\cdots N_n}{n}\right) \xrightarrow{p} g(\lambda)$ ($\therefore \frac{1}{N} \xrightarrow{p} \frac{1}{\lambda}$)

綜合上述事實，得知 $\frac{\sqrt{n}}{N}(\bar{T}-\lambda\mu) \xrightarrow{d} N(0, \frac{1}{\lambda}(\sigma^2+\mu^2))$

㉗

No

Date

$\boxed{12}$ $\begin{pmatrix}\varepsilon\\ v\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \underbrace{\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}}_{\Sigma}\right)$

$W = \alpha S + \varepsilon$ (報酬)

$\therefore \begin{pmatrix}W\\ v\end{pmatrix} = \begin{pmatrix}\varepsilon\\ v\end{pmatrix} + \begin{pmatrix}\alpha S\\ 0\end{pmatrix} \sim N\left(\begin{pmatrix}\alpha S\\ 0\end{pmatrix}, \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right)$

- 將題意解釋為求 $E[W \mid v>0]$.
- 先求 $W \mid v>0$ 之分布

$$\Pr(W \leq w \mid v>0) = \frac{\Pr(W \leq w, v>0)}{\Pr(v>0)}$$

(i) $\Pr(v>0) = \frac{1}{2}$ ($\because$ v 的邊際分布為 $N(0,1)$)

(ii) $\Pr(W \leqq w, v>0) = \Pr(W-\alpha S \leqq w-\alpha S, v>0)$

$= \Pr(\varepsilon \leqq w-\alpha S, v>0)$

→ 需要求複雜的積分，因此先保留。

$$\left(\iint_{\left\{\begin{pmatrix}\varepsilon\\ v\end{pmatrix} \middle| \varepsilon \leqq w-\alpha S, v>0\right\}} \frac{1}{2\pi \cdot |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\varepsilon, v)\Sigma^{-1}\begin{pmatrix}\varepsilon\\ v\end{pmatrix}\right) d\varepsilon dv\right)$$

No.
Date

於是 得到 $\Pr(W \leq w \mid V>0)$

$$f_{W|V>0}(w) = \frac{d}{dw} \Pr(W \leq w \mid V>0)$$

∴ 求 $E[W \mid V>0] = \int_{w\in R} w \cdot \frac{d}{dw} \Pr(W \leq w \mid V>0)\, dw$

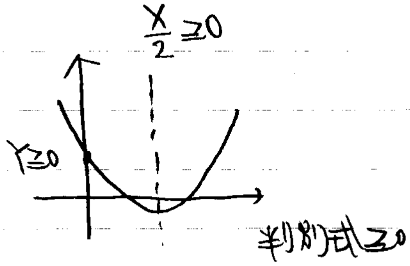
$$= \int_{w\in R} w \cdot \frac{d}{dw} \left\{ \frac{\Pr(\varepsilon \leq w - \alpha S,\ V>0)}{\Pr(V>0)} \right\} dw$$

No.
Date

13 $U, V \overset{iid}{\sim} e(\lambda)$

(a) 考慮二次方程式

$$(t-U)(t-V)=0$$



$$\Leftrightarrow t^2-(U+V)t+UV=0 \quad (\because \text{方程式具有兩個非負根.})$$

$$\Leftrightarrow t^2-Xt+Y=0$$

$$t=\frac{X \pm \sqrt{X^2-4Y}}{2} \quad (X^2-4Y \geq 0,\ X \geq 0,\ Y \geq 0)$$

$$\therefore \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} \frac{X+\sqrt{X^2-4Y}}{2} \\ \frac{X-\sqrt{X^2-4Y}}{2} \end{pmatrix} \text{ or } \begin{pmatrix} \frac{X-\sqrt{X^2-4Y}}{2} \\ \frac{X+\sqrt{X^2-4Y}}{2} \end{pmatrix}$$

$$\text{全機率} = 1 = \iint_{\substack{U \geq 0 \\ V \geq 0}} \lambda^2 \exp(-\lambda(U+V))\, dudv$$

將此積分改為 X、Y 的積分。

case 1: $\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} \frac{X+\sqrt{X^2-4Y}}{2} \\ \frac{X-\sqrt{X^2-4Y}}{2} \end{pmatrix}$ (U≧V 的情況)

$$\frac{\partial}{\partial x}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\left(1+\frac{x}{\sqrt{x^2-4y}}\right) \\ \frac{1}{2}\left(1-\frac{x}{\sqrt{x^2-4y}}\right) \end{pmatrix}$$

$$\frac{1}{2}\left(1+\frac{x}{\sqrt{x^2-4y}}\right)\cdot\frac{1}{\sqrt{x^2-4y}}+\frac{1}{2}\left(1-\frac{x}{\sqrt{x^2-4y}}\right)\cdot\frac{1}{\sqrt{x^2-4y}}$$

No

Date

$$\frac{\partial}{\partial y}\begin{pmatrix}U\\V\end{pmatrix}=\begin{pmatrix}-1/\sqrt{x^2-4y}\\1/\sqrt{x^2-4y}\end{pmatrix}$$

$$\therefore \det J=\frac{1}{2}\left(1+\frac{x}{\sqrt{x^2-4y}}\right)\cdot\frac{1}{\sqrt{x^2-4y}}+\frac{1}{2}\left(1-\frac{x}{\sqrt{x^2-4y}}\right)\cdot\frac{1}{\sqrt{x^2-4y}}$$

$$=\left(\sqrt{x^2-4y}\right)^{-1}$$

$$\therefore\ du\,dv=\left(\sqrt{x^2-4y}\right)^{-1}dx\,dy$$

case2 … $\begin{pmatrix}U\\V\end{pmatrix}=\begin{pmatrix}\frac{X-\sqrt{X^2-4Y}}{2}\\\frac{X+\sqrt{X^2-4Y}}{2}\end{pmatrix}$ (U<Vの時の情形)

同様 $du\,dv=\left(\sqrt{x^2-4y}\right)^{-1}dx\,dy$

$\therefore$ 全機率 $=1=\iint_{\substack{u\geq 0\\ v\geq 0}} f_{UV}(u,v)\,du\,dv$

$$=\iint_{u\geq v>0} f(u,v)\,du\,dv+\iint_{v\geq u>0} f(u,v)\,du\,dv$$

$u\geq v>0 \rightarrow \begin{cases}x^2-4y\geq 0\\ x,y\geq 0\end{cases}$, $f(u,v)=$ case1, $du\,dv \rightarrow \sqrt{x^2-4y}^{\,-1}dx\,dy$

$v\geq u>0 \rightarrow \begin{cases}x^2-4y\geq 0\\ x,y\geq 0\end{cases}$, $f(u,v)=$ case2, $du\,dv \rightarrow \sqrt{x^2-4y}^{\,-1}dx\,dy$

$$=\iint_{\left\{\binom{x}{y}\,\middle|\,\substack{x^2-4y\geq 0\\ x,y\geq 0}\right\}} \lambda^2\exp(-\lambda x)/\sqrt{x^2-4y}\ dx\,dy\times 2.$$

$$\therefore f_{XY}(x,y)=\frac{2\lambda^2\exp(-\lambda x)}{\sqrt{x^2-4y}}\quad\left(\begin{matrix}x,y\geq 0.\\ x^2-4y\geq 0\end{matrix}\right)$$

31

No.

Date

$\boxed{13}$

(b) 接下來求 Y 的 Best Linear Predictor.

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \underset{\binom{\alpha}{\beta}}{\operatorname{argmin}}\ E[(Y-\alpha X-\beta)^2] = S\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

$$S(\alpha,\beta) = E[Y^2+\alpha^2X^2+\beta^2-2\alpha XY+2\alpha\beta X-2\beta Y]$$

$$\begin{cases} ① \ \dfrac{\partial}{\partial\alpha}S\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = 2\alpha E[X^2]-2E[XY]+2\beta E[X]=0 \\ ② \ \dfrac{\partial}{\partial\beta}S\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = 2\beta+2\alpha E[X]-2E[Y]=0 \end{cases}$$

- $① - ② \cdot E[X] = 2\alpha E[X^2]-2E[XY]+\cancel{2\beta E[X]}$

$$\cancel{-2\beta E[X]}-2\alpha E[X]^2+2E[X]E[Y]=0$$

$$\Rightarrow 2\alpha V[X] = 2\,Cov[X,Y]$$

$$\Rightarrow \hat{\alpha} = \frac{Cov[X,Y]}{V[X]}$$

- $② \Leftrightarrow \beta = E[Y]-\alpha E[X] = E[Y]-\dfrac{E[X]}{V[X]}Cov[X,Y]$

$$\Rightarrow \hat{\beta} = E[Y]-\frac{E[X]}{V[X]}Cov[X,Y]$$

Chrys culture

No.

Date

接著求 $E[X]$ $V[X]$. $E[Y]$. $\mathrm{cov}[X,Y]$

(i) $E[X] = \iint_{\substack{X\geq 0,\ Y\geq 0 \\ X^2-4Y\geq 0}} x \cdot \dfrac{2\lambda^2 \exp(-\lambda x)}{\sqrt{x^2-4y}}\, dx\, dy$

先處理 y 的積分：$0 \leq Y \leq \dfrac{x^2}{4}$

$$\int_{y=0}^{y=\frac{x^2}{4}} \lambda^2 \cdot x \cdot \exp(-\lambda x) \frac{1}{\sqrt{\frac{x^2}{4}-y}}\, dy$$

$$= \lambda^2 x \exp(-\lambda x) \cdot \left[(-2)\left(\frac{x^2}{4}-y\right)^{\frac{1}{2}}\right]_{y=0}^{y=\frac{x^2}{4}}$$

$$= \lambda^2 x \exp(-\lambda x) \cdot 2 \cdot \left(\frac{x}{2}\right)$$

$$= \lambda^2 x^2 \exp(-\lambda x)$$

接下來求 $E[X] = \int_{x=0}^{x=\infty} x \cdot \underbrace{\lambda^2 x \exp(-\lambda x)}_{X\text{的邊際 pdf}}\, dx$

$X \sim \Gamma(2, \lambda)$

$\therefore E[X] = \dfrac{2}{\lambda}$

(ii) $\because X \sim \Gamma(2,\lambda)$ $\therefore E[X^2] = \dfrac{6}{\lambda^2}$

$\left(\therefore V[X] = \dfrac{2}{\lambda^2}\right)$

(37)

No.
Date

(iii) $E[Y] = \iint_{\substack{x\geq 0\\ y\geq 0\\ x^2-4y\geq 0}} y \cdot \frac{2\lambda^2 \exp(-\lambda x)}{\sqrt{x^2-4y}} \, dx\, dy$

先處理 y 的積分:

$$\lambda^2 \exp(-\lambda x) \int_{y=0}^{y=\frac{x^2}{4}} \frac{y}{\sqrt{\frac{x^2}{4}-y}} \, dy$$

$$\int_{y=0}^{y=\frac{x^2}{4}} \frac{\left(y-\frac{x^2}{4}\right)+\left(\frac{x^2}{4}\right)}{\sqrt{\frac{x^2}{4}-y}} \, dy$$

$$= \int_{y=0}^{y=\frac{x^2}{4}} \left\{ \frac{x^2}{4} \cdot \frac{1}{\sqrt{\frac{x^2}{4}-y}} - \sqrt{\frac{x^2}{4}-y} \right\} dy$$

$$= \left[-\frac{x^2}{4} \cdot 2 \cdot \left(\frac{x^2}{4}-y\right)^{\frac{1}{2}} + \frac{2}{3}\left(\frac{x^2}{4}-y\right)^{\frac{3}{2}} \right]_{y=0}^{y=\frac{x^2}{4}}$$

$$= \frac{x^2}{4} \cdot \cancel{2} \cdot \left(\frac{x}{\cancel{2}}\right) - \frac{2}{3} \cdot \left(\frac{x}{2}\right)^3$$

$$= \frac{x^3}{4} - \frac{x^3}{12} = \frac{x^3}{6}$$

接著求 $\int_{x=0}^{x=\infty} \lambda^2 \exp(-\lambda x) \cdot \frac{x^3}{6} \, dx$

$z = \lambda x \quad \frac{dz}{dx} = \lambda \quad \int_{z=0}^{z=\infty} \lambda^2 \exp(-z) \cdot \frac{1}{6} \left(\frac{z}{\lambda}\right)^3 \cdot \frac{dz}{\lambda}$

No.
Date

$$= \int_{z=0}^{z=\infty} \frac{1}{6} \cdot z^3 \cdot \frac{1}{\lambda^2} \exp(-z) dz$$

$$= \frac{1}{6\lambda^2} \cdot \Gamma(4) = \frac{1}{\lambda^2} \quad \therefore E[Y] = \frac{1}{\lambda^2}$$

(iv) $E[XY]$ … 可以利用(iii)中間的結果

$$= \int_0^\infty \frac{\lambda^2}{6} x^4 \exp(-\lambda x)\, dx$$

(x的指數 $3 \to 4$)

$z = \lambda x \quad \frac{dz}{dx} = \lambda \qquad \int_0^\infty \frac{\lambda^2}{6} \left(\frac{z}{\lambda}\right)^4 \exp(-z) \cdot \frac{dz}{\lambda}$

$$= \int_0^\infty \frac{1}{6} \frac{1}{\lambda^3} \cdot z^4 \exp(-z) dz$$

$$= \frac{\Gamma(5)}{6\lambda^3} = \frac{4}{\lambda^3}$$

$$\therefore E[XY] = \frac{4}{\lambda^3}$$

由(i)~(iv) $\mathrm{cov}[X,Y] = \frac{4}{\lambda^3} - \frac{2}{\lambda} \cdot \frac{1}{\lambda^2} = \frac{2}{\lambda^3}$

$$\begin{cases} \cdot\ \alpha = \mathrm{cov}[X,Y] / V(X) = \frac{2}{\lambda^3} \Big/ \frac{2}{\lambda^2} = \frac{1}{\lambda} \\ \cdot\ \beta = E[Y] - \frac{E(X)}{V(X)} \mathrm{cov}(X,Y) = \frac{1}{\lambda^2} - \frac{2}{\lambda} \cdot \frac{\lambda^2}{2} \cdot \frac{2}{\lambda^3} = \frac{-1}{\lambda^2} \end{cases}$$

$$\therefore \alpha = \frac{1}{\lambda},\ \beta = \frac{-1}{\lambda^2} \qquad \left(\frac{X}{\lambda} - \frac{1}{\lambda^2}\right)$$

(35)

No

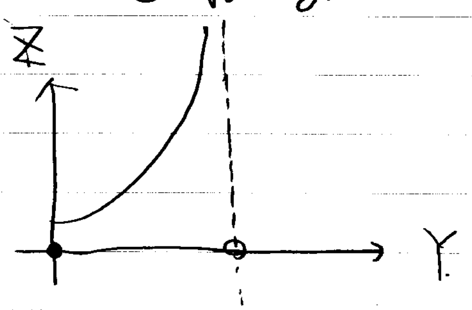
Date

13

(c) 求Y的 Best Predictor Given X

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{2\lambda^2 \exp(-\lambda x)}{\sqrt{x^2-4y}} \div \lambda^2 x \exp(-\lambda x)$$

$$= \frac{2}{x} \cdot \frac{1}{\sqrt{x^2-4y}} \quad \left(0 \le y \le \frac{x^2}{4}\right)$$



$Y = \frac{X^2}{4}$

$f_{Y|X}(y|x)$ 之機率密度函數 離 $\frac{X^2}{4}$ 越近,越大

$\therefore Y = \frac{X^2}{4}$ 為 Best Predictor of Y Given X.

No.
Date

$\boxed{14}$

(a) 狗獾總數… N （這題該考慮有限母體）

- 其中有標誌的個數… K
- 〃 無標誌的個數 $N-K$ $(=M\ ?)$

抽出來的樣本個數… n

觀測到 x 個有標誌的狗獾之機率

$$= \Pr(X=x \mid N, K, n) = \frac{{}_K C_x \cdot {}_{N-K} C_{n-x}}{{}_N C_n}$$

$$\therefore X \sim HG(N, K, n)$$

(b) 根據題意，$K=15$，$n=24$，$x=11$

$$L(N \mid x, K, n) = \frac{{}_K C_x \cdot {}_{N-K} C_{n-x}}{{}_N C_n}$$

$$\hat{N}_{MLE} = \underset{N}{\operatorname{argmax}}\, L(N \mid x, K, n)$$

（用微分求 $\hat{N}_{MLE}$ 應該相當困難，所以考慮 $\frac{L(N)}{L(N-1)}$

- $L(N|x,k,n) = \dfrac{{}_kC_x \cdot {}_{N-k}C_{n-x}}{{}_NC_n}$

$$\frac{L(N|x,k,n)}{L(N-1|x,k,n)} = \frac{\cancel{{}_kC_x} \cdot {}_{N-k}C_{n-x}}{{}_NC_n} \cdot \frac{{}_{N-1}C_n}{\cancel{{}_kC_x} \cdot {}_{N-k-1}C_{n-x}}$$

$$= \underbrace{\frac{{}_{N-1}C_n}{{}_NC_n}}_{①} \cdot \underbrace{\frac{{}_{N-k}C_{n-x}}{{}_{N-k-1}C_{n-x}}}_{②}$$

① … $\dfrac{(N-1)!}{(N-n-1)!\cancel{n!}} \cdot \dfrac{(N-n)!\cancel{n!}}{N!} = \dfrac{N-n}{N}$

② … $\dfrac{(N-k)!}{\cancel{(n-x)!}(N-k-n+x)!} \cdot \dfrac{\cancel{(n-x)!}(N-k-n+x-1)!}{(N-k-1)!}$

$$= \frac{(N-k)}{(N-k-n+x)}$$

$$\therefore \frac{L(N|x,k,n)}{L(N-1|x,k,n)} = ① \times ② = \frac{N-n}{N} \cdot \frac{N-k}{N-k-n+x}$$

考慮 $\dfrac{L(N|x,k,n)}{L(N-1|x,k,n)} \geq 1$ 的範囲

$K=15 \quad n=24 \quad x=1$

(根據題意 $(N-K)-(n-x) \geq 0$)

$\therefore \dfrac{L(N \mid x,k,n)}{L(N-1 \mid x,k,n)} \geq 1$

$\Leftrightarrow \underbrace{(N-n)(N-k)} \geq N(N-k-n+x)$

$N^2-(n+k)N+nk \geq N^2-(n+k)N+Nx$

$\therefore Nx \leq nK$

$\therefore N \leq \dfrac{nK}{x}$

- if $\dfrac{nK}{x} \in$ 整數

$\dfrac{L(2)}{L(1)}>1, \dfrac{L(3)}{L(2)}>1 \cdots \dfrac{L(\frac{nk}{x})}{L(\frac{nk}{x}-1)}=1, \dfrac{L(\frac{nk}{x}+1)}{L(\frac{nk}{x})}<1 \cdots$

$\therefore L(1)<L(2)<\cdots<\underbrace{L(\tfrac{nk}{x}-1)=L(\tfrac{nk}{x})}>\cdots$

$N=\dfrac{nk}{x}$ or $\dfrac{nk}{x}-1$

- if $\dfrac{nk}{x} \notin$ 整數

$\dfrac{L(2)}{L(1)}>1, \dfrac{L(3)}{L(2)}>1 \cdots \dfrac{L([\frac{nk}{x}])}{L([\frac{nk}{x}]-1)}>1, \dfrac{L([\frac{nk}{x}]+1)}{L([\frac{nk}{x}])}<1 \cdots$

(39)

No

Date

$\therefore L(1) < L(2) < \cdots < L\left(\left[\frac{nK}{x}\right]-1\right) < L\left(\left[\frac{nK}{x}\right]\right) > L\left(\left[\frac{nK}{x}\right]+1\right) \cdots$

$\therefore \hat{N}_{MLE} = \left[\frac{nK}{x}\right]$

$x = 11, \quad n = 24, \quad K = 15$

$\because \frac{nK}{x} = \frac{360}{11}$ 非整數

$\therefore \left[\frac{360}{11}\right] = 32$

$\therefore \hat{N}_{MLE} = 32$

(c) $HG(N=32, K=15, n=24)$ $\begin{cases} \text{tagged } 15 \\ \text{untagged } 17 \end{cases}$

($x = 11$ 使得機率最大 ($\because$ (b) $\therefore L(N|x,k,n) = P_r(X|N,k,n)$)

$\underset{n}{24} - \underset{\text{untagged}}{17} = 7 \quad \therefore 7 \leq X \leq 15$)

No

Date

$Pr(X=11 \mid N=32,\ K=15,\ n=24)$

$$= \frac{{}_{15}C_{11} \cdot {}_{17}C_{13}}{{}_{32}C_{24}}$$

(d) 根據(b)的討論，N=32 使得 $L(N \mid X, K, n)$ 最大。

由於 $L(N \mid X, K, n) = Pr(X \mid N, K, n)$，因此

$N = 32$ 時 $Pr(X \mid N, K, n)$ 會低於 (c)：$\frac{{}_{15}C_{11} \cdot {}_{17}C_{13}}{{}_{32}C_{14}}$。

（總而言之 $\hat{N}_{MLE} = 32$. 是 Maximum Likelihood Estimator，

無論是 機率 還是 概似函數，N=32 時 會最大）

R05246013

No.
Date

高等統計推論 II 作業3. 森元俊成

$\boxed{1}$ $f(x|\theta) = \frac{1}{\pi} \cdot \frac{1}{1+(x-\theta)^2}$ $(-\infty < x < \infty,\ \theta \in \mathbb{R})$

$$f(x_1,x_2,x_3|\theta) = \frac{1}{\pi^3} \cdot \frac{1}{1+(x_1-\theta)^2} \cdot \frac{1}{1+(x_2-\theta)^2} \cdot \frac{1}{1+(x_3-\theta)^2}$$

$X_1 = 0,\ X_2 = 1,\ X_3 = a$

$$\therefore f(0,1,a|\theta) = \frac{1}{\pi^3} \cdot \frac{1}{1+\theta^2} \cdot \frac{1}{1+(1-\theta)^2} \cdot \frac{1}{1+(a-\theta)^2}$$

$$\therefore L(\theta) \overset{\text{def}}{=} \frac{1}{\pi^3} \cdot \frac{1}{1+\theta^2} \cdot \frac{1}{1+(1-\theta)^2} \cdot \frac{1}{1+(a-\theta)^2} \quad (\theta \in \mathbb{R})$$

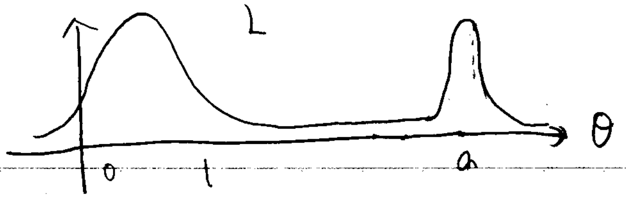
$$l(\theta) = \log L(\theta) = -3\log\pi - \log(1+\theta^2) - \log(1+(1-\theta)^2) - \log(1+(a-\theta)^2)$$

$$\underline{\underset{||}{l'(\theta)}} = \frac{-2\theta}{1+\theta^2} + \frac{2(1-\theta)}{1+(1-\theta)^2} + \frac{2(a-\theta)}{1+(a-\theta)^2}$$

$\frac{\partial l(\theta)}{\partial\theta}$

- $l'(0) = 2 + \frac{2a}{1+a^2} \approx 2$ $\left(\because a \text{很大},\ \therefore \frac{2a}{1+a^2} \approx 0\right)$
- $l'(1) = -1 + \frac{2(a-1)}{1+(a-1)^2} \approx -1$ $\left(\because a \text{很大}\quad \frac{2(a-1)}{1+(a-1)^2} \approx 0\right)$
- $l'(a) = \frac{-2a}{1+a^2} + \frac{2(1-a)}{1+(1-a)^2} \approx 0$ (但<0)

(圖)



① 由於 $l'(0)\approx 2>0$ 且 $l'(1)\approx -1<0$,

故 $(0,1)$ 之間一定存在 $l'(x)$ 從「正」變成「負」的點

∴ 區間 $(0,1)$ 之間發生極大值

② $l'(1)\approx -1$, $l'(a)\approx 0$, $l'(a)<0$

但考慮很小的正數 ε,例如… $\varepsilon=\frac{100}{a}$ (假 $a=100000\ldots$)

$$l'(a-\varepsilon)=\frac{-2a}{1+(a-\varepsilon)^2}+\frac{2(1-a+\varepsilon)}{1+(1-a+\varepsilon)^2}+\frac{2\varepsilon}{1+\varepsilon^2}$$

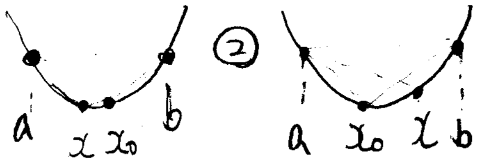
$$\approx -\frac{2}{a} \qquad \approx -\frac{2}{a} \qquad \frac{\frac{200}{a}}{1+\frac{10000}{a^2}}=\frac{200a}{10000+a^2}\approx\frac{200}{a}$$

$$\therefore\ l'(a-\varepsilon)\approx -\frac{2}{a}-\frac{2}{a}+\frac{200}{a}=\frac{196}{a}>0$$

由此可知,比 a 稍微小一點的地方存在 $l'(x)>0$
($x\approx a$,但是 $x<a$)

故區間 $(a-\varepsilon,a)$ 之間也存在 $l'(x)$ 從「正」變成「負」的地方,∴ $(a-\varepsilon,a)\subseteq(1,a)$ (∴也可以說 $(1,a)$ 之間) 發生極大值(實際上是離 a 很近的地方)

③ (a)… ① ②



$\boxed{2}$ $f(x)\ (x \in C \subseteq \mathbb{R})$

(a) 請參閱「後面 Note $\boxed{2}$」

「$\dfrac{f(y)-f(x)}{y-x} \leq \dfrac{f(z)-f(x)}{z-x} \leq \dfrac{f(z)-f(y)}{z-y}$」 $(x \leq y \leq z)$

考慮開區間 (a,b) 上以及 $x, x_0 \in (a,b)$

① $x<x_0$ $\dfrac{f(x)-f(a)}{x-a} \leq \dfrac{f(x_0)-f(a)}{x_0-a} \leq \dfrac{f(x_0)-f(x)}{x_0-x} \leq \dfrac{f(b)-f(x)}{b-x} \leq \dfrac{f(b)-f(x_0)}{b-x_0}$

or

② $x_0<x$ $\dfrac{f(x_0)-f(a)}{x_0-a} \leq \dfrac{f(x)-f(a)}{x-a} \leq \dfrac{f(x)-f(x_0)}{x-x_0} \leq \dfrac{f(b)-f(x_0)}{b-x_0}$

無論 ① or ②, $\left|\dfrac{f(x)-f(x_0)}{x-x_0}\right| \leq \max\left\{\left|\dfrac{f(b)-f(x_0)}{b-x_0}\right|, \left|\dfrac{f(x_0)-f(a)}{x_0-a}\right|\right\}$

$\therefore |f(x)-f(x_0)| \leq |x-x_0| \max\left\{\left|\dfrac{f(b)-f(x_0)}{b-x_0}\right|, \left|\dfrac{f(x_0)-f(a)}{x_0-a}\right|\right\}$

$\therefore x \to x_0 \Rightarrow |f(x)-f(x_0)| \to 0$ 「註 a, b, x_0 … fixed」

$\therefore f$ 於 $\forall x_0 \in (a,b)$ 連續.

(b) Note $\boxed{2}$ ① 「$\dfrac{f(y)-f(x)}{y-x} \leq \dfrac{f(z)-f(x)}{z-x}$」 $(x \leq y \leq z)$

(i) $y \downarrow x \Rightarrow$ 得 $f'(x) \leq \dfrac{f(z)-f(x)}{z-x}$
(∵可微)

No

Date

(iii) $z \to y \Rightarrow$ 得 $f'(x) \leq \dfrac{f(y)-f(x)}{y-x}$ ∴ $f'(x)(y-x) \leq f(y)-f(x)$ ($x \leq y$ 時成立)

· 同樣道理 ② $\dfrac{f(z)-f(x)}{z-x} \leq \dfrac{f(z)-f(y)}{z-y}$ … $(x<y\leq z)$

$z \to y \Rightarrow$ 得 $\dfrac{f(y)-f(x)}{y-x} \leq f'(y)$ ∴ $f(y)-f(x) \leq (y-x)f'(y)$
$(x \leq y)$

$\Rightarrow$ 替換 x 與 y 「$f(x)-f(y) \leq (x-y)f'(x)$ $(y \leq x)$」

$\Leftrightarrow$ 「$f'(x)(y-x) \leq f(y)-f(x)$ $(y \leq x)$」

由此可知 無論 $x \leq y$ or $y \leq x$ 皆成立

(c) 利用(b)的結果…

① $y \geq x$ …「$f'(x) \leq \dfrac{f(y)-f(x)}{y-x}$」… $y \searrow x$ 得「$f'(x) \leq f'(x+0)$」

② $y \leq x$ …「$\dfrac{f(x)-f(y)}{x-y} \leq f'(x)$」… $y \nearrow x$ 得「$f'(x-0) \leq f'(x)$」

由①,②得知, $f'(x-0) \leq f'(x) \leq f'(x+0)$

∴ f' 為非減函數, $f''(x) \geq 0$

⑤

No.
Date

$\boxed{3}$

$f(x_1, x_2, \ldots x_n | \lambda) = \lambda^n \exp(-\lambda(x_1+x_2+\ldots x_n))$ $(x_i \geq 0)$

$\log f(x_1 \ldots x_n | \lambda) = n \log \lambda - \lambda(X_1+X_2+\ldots+X_n)$

$\frac{\partial}{\partial \lambda} \log f(X_1 \sim X_n | \lambda) = \frac{n}{\lambda} - (X_1+X_2+\ldots+X_n) = 0 \Rightarrow \lambda = \frac{n}{X_1+X_2+\ldots+X_n}$

$\frac{\partial}{\partial \lambda^2} \log f(X_1 \sim X_n | \lambda) = \frac{-n}{\lambda^2} < 0$

$\because L(\lambda | X_1 \ldots X_n) = f(x_1 \ldots x_n | \lambda)$

$\lambda = \frac{1}{\overline{X}}$ 使得 $L(\lambda | X_1 \sim X_n)$ 為最大

$\therefore \hat{\lambda}_{MLE} = \frac{1}{(\overline{X})}$

(a) $T \stackrel{def}{=} X_1+X_2+\ldots+X_n \sim \Gamma(n, \lambda)$

$E\left[\frac{1}{T}\right] \left(= E\left[\frac{1}{X_1+X_2+\ldots+X_n}\right]\right)$

$= \int_{T=0}^{T=\infty} \frac{\lambda^n t^{n-2}}{\Gamma(n)} \exp(-\lambda t)\, dt \qquad x \stackrel{let}{=} \lambda t \quad \frac{dx}{dt} = \lambda$

$= \int_0^\infty \frac{\lambda^n}{\Gamma(n)} \left(\frac{x}{\lambda}\right)^{n-2} \exp(-x) \cdot \frac{dx}{\lambda} = \frac{\lambda}{\Gamma(n)} \cdot \Gamma(n-1) = \frac{\lambda}{n-1}$

No.
Date

$\therefore E\left[\frac{n}{T}\right]=\frac{n}{n-1}\lambda$

$\therefore E\left[\frac{1}{\bar{X}}\right]=\frac{n}{n-1}\lambda\neq\lambda$ $\therefore$ $\hat{\lambda}_{MLE}$並非不偏估計量

(b) $E[X_j]=\underset{(<\infty)}{\frac{1}{\lambda}}$ $(j=1\sim n)$, $X_1\sim X_n$: i.i.d.

根據弱大數法則 (Khinchine),

$\frac{X_1+X_2+\cdots+X_n}{n}\xrightarrow{P}\frac{1}{\lambda}$.

$g(x)=\frac{1}{x}$ $(x\neq 0)$ 為連續函數

$\therefore g\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)\xrightarrow{P}g\left(\frac{1}{\lambda}\right)=\lambda$

$\therefore \frac{1}{\bar{X}}\xrightarrow{P}\lambda$ $\therefore$ $\hat{\lambda}_{MLE}$為一致估計量.

(c) $V[X_j]=\underset{(<\infty)}{\frac{1}{\lambda^2}}$ $(j=1\sim n)$ $X_1\sim X_n$: i.i.d.

根據中央極限定理, $\sqrt{n}\left(\frac{X_1+X_2+\cdots+X_n}{n}-\frac{1}{\lambda}\right)\xrightarrow{d}N\left(0,\frac{1}{\lambda^2}\right)$.

利用 δ-method: $g(x)=\frac{1}{x}$, $g(\bar{X})\approx g'\left(\frac{1}{\lambda}\right)\left(\bar{X}-\frac{1}{\lambda}\right)+g\left(\frac{1}{\lambda}\right)$

$\therefore\sqrt{n}\left(g(\bar{X})-g\left(\frac{1}{\lambda}\right)\right)\approx\underbrace{g'\left(\frac{1}{\lambda}\right)}_{(-\lambda^2)}\cdot\underbrace{\sqrt{n}\left(\bar{X}-\frac{1}{\lambda}\right)}_{N\left(0,\frac{1}{\lambda^2}\right)}\xrightarrow{d}N(0,\lambda^2)$

$\therefore\sqrt{n}\left(\hat{\lambda}_{MLE}-\lambda\right)\xrightarrow{d}N(0,\lambda^2)$

Chuyu culture

⑦

No.
Date

$\boxed{4}$ 敘述利用 Gradient Descent 求 MLE 的方法

$$f(x_1, x_2, \cdots, x_n \mid \alpha, \beta) = \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} (x_1 x_2 \cdots x_n)^{\alpha-1} \exp(-\beta(x_1 + \cdots + x_n))$$

$$(x_1 \sim x_n \geq 0)$$

$$\theta \overset{def}{=} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$$L(\theta) = L\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \overset{def}{=} \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} (x_1 x_2 \cdots x_n)^{\alpha-1} \exp(-\beta(x_1 + x_2 + \cdots + x_n))$$

$$(\alpha > 0, \beta > 0)$$

「將 $L(\theta)$ 最大化」等於「將 $\log L(\theta)$ 最大化」

在此考慮 $\log L(\theta \mid \vec{x})$ 之最大化.

$$\log L(\theta \mid \vec{x}) = n\alpha \log \beta - n \log \Gamma(\alpha) + (\alpha - 1) \sum_{k=1}^{n} \log X_k - \beta \sum_{k=1}^{n} X_k$$

$$l(\theta) \left(= l\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right) \overset{def}{=} \log L(\theta \mid \vec{x})$$

$$\frac{\partial l}{\partial \theta} = \begin{pmatrix} \frac{\partial l}{\partial \alpha} \\ \frac{\partial l}{\partial \beta} \end{pmatrix} = \begin{pmatrix} n \log \beta - \frac{\partial}{\partial \alpha} n \log \Gamma(\alpha) + \sum_{k=1}^{n} \log X_k \\ \frac{n\alpha}{\beta} - \sum_{k=1}^{n} X_k \end{pmatrix}$$

㊟ $\Gamma(\alpha)$ 之 α 不一定為整數, 故其微分直接寫成

$$\frac{\partial \Gamma(\alpha)}{\partial \alpha} = \frac{\partial}{\partial \alpha} \int_0^\infty x^{\alpha-1} \exp(-x)\, dx$$

No.
Date

- step 0 … 設定初始值 $\hat{\theta}_0 = \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}$ $(\alpha_0, \beta_0 > 0)$
- step j … 依照以下公式更新 $\hat{\theta}_{j-1} \to \hat{\theta}_j$ $(j \geq 1)$

$$\hat{\theta}_j \overset{\text{update}}{=} \hat{\theta}_{j-1} - c \cdot \frac{\partial l}{\partial \theta}\bigg|_{\theta = \binom{\alpha}{\beta} = \hat{\theta}_{j-1}}$$

where $\hat{\theta}_j = \begin{pmatrix} \hat{\alpha}_j \\ \hat{\beta}_j \end{pmatrix}$, $\frac{\partial l}{\partial \theta} = \begin{pmatrix} n \log \beta - \frac{\partial}{\partial \alpha} n \log \Gamma(\alpha) + \sum_{k=1}^{n} \log X_k \\ \frac{n\alpha}{\beta} - \sum_{k=1}^{n} X_k \end{pmatrix}$

c … 常數

- $\hat{\theta}_j$ 收斂時 ($\hat{\theta}_j - \hat{\theta}_{j-1} \approx 0$ 時) 停止計算，

並將 $\hat{\theta}_j$ 作為 $\hat{\theta}_{MLE} = \begin{pmatrix} \alpha_{MLE} \\ \beta_{MLE} \end{pmatrix}$

Chryo culture

No.
Date

$\boxed{5}$ $\sigma \xrightarrow{\text{改用}} \beta$. $f(x|\beta) = \frac{1}{2\beta} \exp\left(-\frac{1}{\beta}|x|\right)$

① Method of Moment Estimator 與其一致性:

$E[X] = 0$,

$E[X^2] = \int_{-\infty}^{\infty} \frac{x^2}{2\beta} \cdot \exp\left(-\frac{1}{\beta}|x|\right) dx = \int_0^{\infty} \frac{x^2}{\beta} \exp\left(-\frac{1}{\beta}x\right) dx$

$\left(y \overset{\text{put}}{=} \frac{x}{\beta} \quad \frac{dy}{dx} = \frac{1}{\beta} \right) = \int_0^{\infty} \beta y^2 \exp(-y) \cdot \beta dy$

$= \int_0^{\infty} \beta^2 y^2 \exp(-y) dy = \beta^2 \cdot \Gamma(3) = 2\beta^2 < \infty$

根據弱大數法則, $\frac{X_1^2 + \cdots X_n^2}{n} \xrightarrow{P} 2\beta^2$

$g_0(x) \overset{\text{def}}{=} \sqrt{\frac{x}{2}}$ $(x > 0)$ 為連續函數,

$\therefore g_0\left(\frac{X_1^2 + X_2^2 + \cdots + X_n^2}{n}\right) \xrightarrow{P} g_0(2\beta^2) = \beta \quad (\because \beta > 0)$

$\hat{\beta}_{MME} = \sqrt{\frac{X_1^2 + X_2^2 + \cdots + X_n^2}{2n}} \quad (\hat{\beta}_{MLE} \xrightarrow{P} \beta)$

接下來考慮 $\hat{\beta}_{MME}$ 之極限分布. ($\sqrt{n}(\hat{\beta}_{MME} - \beta)$ 之極限分布)

$W_i \overset{\text{def}}{=} \frac{X_i^2}{2}$ $(i = 1 \sim n)$ $\left(\hat{\beta}_{MME} = \sqrt{\frac{W_1 + W_2 + \cdots + W_n}{n}} \right)$

$E[W_j] = \beta^2 \quad E[W_j^2] = E[\frac{x^4}{4}] = \int_{-\infty}^{\infty} \frac{x^4}{4} \cdot \frac{1}{2\beta} \exp(\frac{-1}{\beta}|x|) dx$

$= \int_0^{\infty} \frac{x^4}{4} \cdot \frac{1}{\beta} \exp(\frac{-1}{\beta}x) dx \qquad y = \frac{1}{\beta}x \qquad \frac{dy}{dx} = \frac{1}{\beta}$

$= \int_0^{\infty} \frac{1}{4} \cdot (\beta y)^4 \cdot \frac{1}{\beta} \cdot \exp(-y)\, \beta dy$

$= \int_0^{\infty} \frac{\beta^4}{4} y^4 \exp(-y) dy = \frac{\beta^4}{4} \cdot \Gamma(5) = 6\beta^4$

$\therefore V[W_j] = E[W_j^2] - E[W_j]^2 = 5\beta^4 < \infty$

根據中央極限定理，$\sqrt{n}\left(\frac{W_1+W_2+\cdots+W_n}{n} - \beta^2\right) \xrightarrow{d} N(0, 5\beta^4)$

利用 δ-method. $g(x) \overset{def.}{=} \sqrt{x} \qquad g'(x) = \frac{1}{2\sqrt{x}}$

$\sqrt{n}\left(\underbrace{g\left(\frac{W_1+W_2+\cdots+W_n}{n}\right)}_{= \hat{\beta}_{MME}} - g(\beta^2)\right) \xrightarrow{d} N(0, \underbrace{(g'(\beta^2))^2 \cdot 5\beta^4}_{})$

$= N(0, \frac{5}{4}\beta^2)$

$\therefore \sqrt{n}(\hat{\beta}_{MME} - \beta) \xrightarrow{d} N(0, \frac{5}{4}\beta^2)$

$\boxed{5}$ - part 2

② MLE與其一致性…。

$$f(x_1,\dots,x_n|\beta)=\left(\frac{1}{2\beta}\right)^n \exp\left(-\frac{1}{\beta}(|x_1|+\dots+|x_n|)\right) \quad (\beta>0)$$

$$\log f(x_1,\dots,x_n|\beta)=-n\log(2\beta)-\frac{1}{\beta}\{|x_1|+|x_2|+\dots+|x_n|\}$$

$$\frac{\partial}{\partial\beta}\log f(x_1,\dots,x_n|\beta)=-\frac{n}{\beta}+\frac{1}{\beta^2}\{|x_1|+\dots|x_n|\}=\frac{|x_1|+\dots|x_n|-n\beta}{\beta^2}$$

$$\Rightarrow \beta=\frac{1}{n}(|x_1|+\dots+|x_n|)$$

β	0		$\frac{1}{n}(\lvert x_1\rvert+\dots\lvert x_n\rvert)$	
$\frac{\partial}{\partial\beta}\log f$		$+$	0	$-$
$\log f$		$\nearrow$	max	$\searrow$

$$\therefore \beta_{MLE}=\frac{1}{n}\sum_{j=1}^{n}|X_j|$$

求 $|X_j|$ 之分佈…　$Z=|X|$

$$\text{全機率}=1=\int_{-\infty}^{\infty}\left(\frac{1}{2\beta}\right)\exp\left(-\frac{1}{\beta}|x|\right)dx$$

$$=\int_0^\infty \frac{1}{2\beta}\exp\left(-\frac{1}{\beta}|x|\right)dx+\int_{-\infty}^{0}\frac{1}{2\beta}\exp\left(-\frac{1}{\beta}|x|\right)dx$$

$0\to\infty$　$|x| \to z$　$dx \to dz$　$-\infty\to 0$　$|x| \to z$　$dx \to -dz$

No.
Date

$$= \int_0^\infty \frac{1}{2\beta} \exp\left(\frac{-1}{\beta} z\right) dz + \int_\infty^0 \frac{1}{2\beta} \exp\left(\frac{-1}{\beta} z\right)(-dz)$$

$$= \int_0^\infty \frac{1}{\beta} \exp\left(\frac{-1}{\beta} z\right) dz \quad \therefore Z = |X| \sim \exp\left(\frac{1}{\beta}\right)$$

$$\hat{\beta}_{MLE} = \frac{1}{n}(Z_1 + Z_2 + \dots Z_n) \quad \left(Z_j \sim \exp\left(\frac{1}{\beta}\right)\right)$$

(註 根據弱大數法則, $\hat{\beta}_{MLE} \xrightarrow{p} \beta$)

$$V[Z_j] = \beta^2 < \infty$$

∴ 根據中央極限定理, $\sqrt{n}\left(\frac{Z_1+Z_2+\dots Z_n}{n} - \beta\right) \xrightarrow{d} N(0, \beta^2)$

$$\therefore \sqrt{n}(\hat{\beta}_{MLE} - \beta) \xrightarrow{d} N(0, \beta^2)$$

$\boxed{6}$.

X	Pr
-2	$\frac{1-\theta}{4}$
-1	$\frac{\theta}{12}$
0	$\frac{1}{2}$
1	$\frac{3-\theta}{12}$
2	$\frac{\theta}{4}$

(a)

$$E[X] = (-2)\cdot\frac{1-\theta}{4} + (-1)\cdot\frac{\theta}{12} + 0\cdot\left(\frac{1}{2}\right) + (1)\cdot\frac{3-\theta}{12} + (2)\left(\frac{\theta}{4}\right)$$

$$= \frac{\theta-1}{2} - \frac{\theta}{12} + \frac{3-\theta}{12} + \frac{\theta}{2}$$

$$= \frac{6\theta-6-\theta+3-\theta+6\theta}{12} = \frac{10\theta-3}{12}$$

$$E\left[\frac{12X}{10}\right] = \theta - 0.3 \quad \therefore E[1.2X+0.3] = \theta$$

$\therefore 1.2X+0.3$ 為 θ 之不偏估計量

(b)

$$Pr(X|\theta) = \left(\frac{1-\theta}{4}\right)^{I(X=-2)} \cdot \left(\frac{\theta}{12}\right)^{I(X=-1)} \cdot \left(\frac{1}{2}\right)^{I(X=0)} \cdot \left(\frac{3-\theta}{12}\right)^{I(X=1)} \cdot \left(\frac{\theta}{4}\right)^{I(X=2)}$$

$$\therefore L(\theta|X=x) = \left(\frac{1-\theta}{4}\right)^{I(X=-2)} \cdot \left(\frac{\theta}{12}\right)^{I(X=-1)} \left(\frac{1}{2}\right)^{I(X=0)} \cdot \left(\frac{3-\theta}{12}\right)^{I(X=1)} \left(\frac{\theta}{4}\right)^{I(X=2)}$$

$(0 \leq \theta \leq 1)$

考慮θ使得$L(\theta|x)$為最大.

$L(\theta|X=-2)=\frac{1-\theta}{4}$ ∴ $\theta=0$時最大 $\hat{\theta}_{MLE}=0$

$L(\theta|X=-1)=\frac{\theta}{12}$ ∴ $\theta=1$時最大 $\hat{\theta}_{MLE}=1$

$L(\theta|X=0)=\frac{1}{2}$ ∴ $\forall c\in[0,1)$ $\hat{\theta}_{MLE}=c$

$L(\theta|X=1)=\frac{3-\theta}{12}$ ∴ $\theta=0$時最大 $\hat{\theta}_{MLE}=0$

$L(\theta|X=2)=\frac{\theta}{4}$ ∴ $\theta=1$時最大 $\hat{\theta}_{MLE}=1$

由此可知, $\hat{\theta}_{MLE}$並不唯一 (觀測到$X=0$時)

$$E[\hat{\theta}_{MLE}]=\sum_{x\in\{-2,-1,0,1,2\}} Pr(X=x)\cdot\hat{\theta}_{MLE}|x$$

$$=\frac{1-\theta}{4}\cdot 0+\frac{\theta}{12}\cdot 1+\frac{1}{2}\cdot c+\frac{3-\theta}{12}\cdot 0+\frac{\theta}{4}\cdot 1$$

$$=\frac{\theta}{12}+\frac{c}{2}+\frac{\theta}{4}=\frac{\theta}{3}+\frac{c}{2}$$

無論$c\in[0,1)$為多少, 由於c與θ無關,

故無法使得$E[\hat{\theta}_{MLE}]=\theta$

∴ $\hat{\theta}_{MLE}$並非不偏估計量

15

Date

7

(a) 先求 $X_{(1)}$ 與 $X_{(n)}$ 之聯合分布:

$$\Pr(X_{(1)}\leq x,\ X_{(n)}\leq y) = \Pr(X_{(n)}\leq y) - \Pr(x< X_{(1)}, X_{(n)}\leq y)$$

$$= \left(\frac{y-\theta+1}{2}\right)^n - \left(\frac{y-x}{2}\right)^n \quad \left((x,y)\subseteq(\theta-1,\theta+1)\right)$$

$$\frac{\partial^2}{\partial x\partial y}\Pr(X_{(1)}\leq x,\ X_{(n)}\leq y) = \frac{n(n-1)(y-x)^{n-2}}{2^n}$$

$$\therefore f_{X_{(1)},X_{(n)}}(x,y) = \frac{n(n-1)}{2^n}(y-x)^{n-2} \quad (\theta-1\leq x\leq y\leq \theta+1)$$

$$\begin{cases} r = y-x \\ s = \frac{y+x}{2} \end{cases} \Leftrightarrow \begin{cases} x = s-\frac{r}{2} \\ y = s+\frac{r}{2} \end{cases}$$

$$(dr\,ds = dx\,dy)$$

$$\theta-1 \leq s-\frac{r}{2} \leq s+\frac{r}{2} \leq \theta+1$$

$$\therefore 0\leq \frac{r}{2}\leq \min\{s-\theta+1,\ \theta+1-s\}$$

$$1 = \iint_{\theta-1\leq x\leq y\leq\theta+1} \frac{n(n-1)}{2^n}(y-x)^{n-2}\,dx\,dy$$

$$= \iint_{\theta-1\leq s-\frac{r}{2}\leq s+\frac{r}{2}\leq\theta+1} \frac{n(n-1)}{2^n} r^{n-2}\,dr\,ds$$

culture

No.
Date

$$f_S(s) = \int_0^{2\min\{s-\theta+1,\ \theta+1-s\}} \frac{n(n-1)}{2^n} r^{n-2}\, dr$$

$$= \frac{n}{2^n}\left[r^{n-1}\right]_0^{2\min\{s-\theta+1,\ \theta+1-s\}}$$

$$= \frac{n}{2}\min\{s-\theta+1,\ \theta+1-s\}^{n-1}$$

$$\begin{cases} \cdot\ \theta-1 \leq s \leq \theta \quad f_S(s) = \frac{n}{2}(s-\theta+1)^{n-1} & \cdots \left(\frac{X_{(1)}+X_{(n)}}{2} \text{ 之 pdf}\right) \\ \cdot\ \theta < s \leq \theta+1 \quad f_S(s) = \frac{n}{2}(\theta+1-s)^{n-1} \end{cases}$$

• 接著證明 $S_n \xrightarrow{a.s} \theta$

$\quad\ \ \overset{\shortparallel}{S}$

※ 為了方便，先求 $Pr(|S_n-\theta|<\varepsilon) = \int_{\theta-\varepsilon}^{\theta} f_S(s)\,ds + \int_{\theta}^{\theta+\varepsilon} f_S(s)\,ds$

$$= \int_{\theta-\varepsilon}^{\theta} \frac{n}{2}(s-\theta+1)^{n-1}\,ds + \int_{\theta}^{\theta+\varepsilon} \frac{n}{2}(\theta+1-s)^{n-1}\,ds$$

$$= \left[\frac{1}{2}(s-\theta+1)^n\right]_{\theta-\varepsilon}^{\theta} + \left[\frac{-1}{2}(\theta+1-s)^n\right]_{\theta}^{\theta+\varepsilon}$$

$$= \frac{1-(1-\varepsilon)^n}{2} + \frac{1-(1-\varepsilon)^n}{2} = \underline{1-(1-\varepsilon)^n}$$

$$\begin{cases} \Rightarrow Pr(|S_n-\theta|<\varepsilon) = 1-(1-\varepsilon)^n \\ Pr(|S_n-\theta|\geq\varepsilon) = (1-\varepsilon)^n \end{cases}$$

[7] - part2.

※ $S_n \xrightarrow{a.s} \theta \Leftrightarrow \Pr(\lim_{n\to\infty} S_n=\theta)=1$

$\Leftrightarrow \Pr(\limsup_n |S_n-\theta| \geq \varepsilon)=0$

$\Leftrightarrow \Pr(\underbrace{\lim_n \sup_{m\geq n}}_{\hookrightarrow \text{遞減}} |S_m-\theta| \geq \varepsilon)=0 = \Pr(\bigcap_{n=1}^{\infty} \underbrace{\{\sup_{m\geq n} |S_m-\theta| \geq \varepsilon\}}_{A_n})$

$\Leftrightarrow \lim_{n\to\infty} \Pr(A_n)=0$ $(\because \Pr(A_1)<\infty)$

$\Leftrightarrow \lim_{n\to\infty} \Pr(\bigcup_{m\geq n}^{\infty} \{|S_m-\theta| \geq \varepsilon\})=0$

註 $(\because A_n \supseteq A_{n+1},\quad \Pr(A_1)<\infty$

$\therefore \lim_{n\to\infty} \Pr(A_n) = \Pr(\bigcap_{n} A_n)\ )$

總之, 證明:

$\lim_{n\to\infty} \Pr(\bigcup_{m\geq n}^{\infty} \{|S_m-\theta| \geq \varepsilon\})=0$

$0 \leq \Pr(\bigcup_{m\geq n}^{\infty} \{|S_m-\theta| \geq \varepsilon\}) \leq \sum_{m\geq n}^{\infty} \Pr(|S_m-\theta| \geq \varepsilon)$ (測度之 sub-additivity)

$= \sum_{m\geq n}^{\infty} (1-\varepsilon)^m = (1-\varepsilon)^n \cdot \frac{1}{\varepsilon}$

$0 \leq \lim_{n\to\infty} \Pr(\bigcup_{m\geq n}^{\infty} \{|S_m-\theta| \geq \varepsilon\}) \leq \lim_{n\to\infty} \frac{1}{\varepsilon}(1-\varepsilon)^n = 0$ (for all $\varepsilon \in (0,1)$)

$\therefore$ 由此可知 $S_n \xrightarrow{a.s} \theta$

(b) 證明 $\lim_{n\to\infty} E\left[|\overline{X_n}-\theta|^2\right]=0$

由於 $E[\overline{X_n}]=\theta$ $\left(\because E\left[\frac{X_{(1)}+\dots X_{(n)}}{n}\right]=E\left[\frac{X_1+X_2+\dots X_n}{n}\right]=\theta\right)$

故 $E\left[(\overline{X_n}-\theta)^2\right]=V[\overline{X_n}]=V\left[\frac{X_1+X_2+\dots X_n}{n}\right]$

$=\frac{1}{n^2}V[X_1+X_2+\dots X_n]=\frac{1}{n}V[X_j]$ $(\because X_1\sim X_n : \text{iid})$

$=\frac{1}{n}\frac{1}{12}2^2=\frac{1}{3n}$

$\therefore \lim_{n\to\infty} E\left[(\overline{X_n}-\theta)^2\right]=\lim_{n\to\infty}\frac{1}{n}V(X_j)=\lim_{n\to\infty}\frac{1}{3n}=0$

$\therefore \overline{X_n}\xrightarrow{L_2}\theta$

註 題目中 $X_{(1)}+X_{(2)}+\dots+X_{(n)} = X_1+X_2+\dots+X_n$

(不需考慮順序統計量)

$\boxed{8}$ (a) $\log f(x|\theta) = \log h(x) + \eta(\theta)T(x) - A(\theta)$

$$\frac{\partial}{\partial\theta}\log f(x|\theta) = \eta'(\theta)T(x) - A'(\theta) = 0$$

㊀ ∵ $\hat\theta$ 滿足 $\eta'(\hat\theta)T(x) - A'(\hat\theta) = 0$ （$\hat\theta$ 為 x 之函數）

$$\begin{cases} \log l(\hat\theta) = \log h(x) + \eta(\hat\theta)T(x) - A(\hat\theta) \\ \log l(\theta_0) = \log h(x) + \eta(\theta_0)T(x) - A(\theta_0) \end{cases}$$

$$\log l(\hat\theta) - \log l(\theta_0) = \big(\eta(\hat\theta) - \eta(\theta_0)\big)T(x) - A(\hat\theta) + A(\theta_0)$$

$$\frac{\partial}{\partial x}\big(\log l(\hat\theta) - \log l(\theta_0)\big) = \eta'(\hat\theta)\hat\theta'(x)T(x) + \big(\eta(\hat\theta) - \eta(\theta_0)\big)T'(x) - A'(\hat\theta)\cdot\hat\theta'(x)$$

$$= \underbrace{\hat\theta'(x)\big(\eta'(\hat\theta)T(x) - A'(\hat\theta)\big)}_{=0\ (\because ㊀)} + \big(\eta(\hat\theta) - \eta(\theta_0)\big)T'(x)$$

$$= \underbrace{\big(\eta(\hat\theta) - \eta(\theta_0)\big)}_{>0}T'(x)$$

這題應該需要假設
要 $\hat\theta(x) \geq \theta_0 \Rightarrow T'(x) \geq 0$

$$\begin{cases} \eta：遞增 \\ \because \hat\theta > \theta_0 \Rightarrow \underbrace{\eta(\hat\theta) - \eta(\theta_0)}_{>0} \end{cases}$$

若此假設成立

$\Rightarrow \frac{\partial}{\partial x}\big(\log l(\hat\theta) - \log l(\theta_0)\big) \geq 0$

∴ 證明完成

(b) 題目應為「 $A(\theta) = \log \int h(x) \exp(\eta(\theta) T(x)) dx$ 」 show

由於 $f(x|\theta)$ 為 pdf, 故 $\int_{x\in R} h(x) \exp(\eta(\theta) T(x) - A(\theta)) dx = 1$.

$\therefore e^{A(\theta)} \int_{x\in R} h(x) \exp(\eta(\theta) T(x) - A(\theta)) dx = e^{A(\theta)}$

$= \int_{x\in R} h(x) \exp(\eta(\theta) T(x)) dx = e^{A(\theta)}$

取兩邊的 log $\rightarrow$ $\log \int_{x\in R} h(x) \exp(\eta(\theta) T(x)) dx = A(\theta)$

$\therefore$ 證明完成 ∎

②

$\boxed{9}$ $f(x_{ij}\mid\mu_i,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x_{ij}-\mu_i)^2\right)$

$f(x_{11}\sim x_{1k},\dots,x_{n1}\sim x_{nk}\mid\mu_1\sim\mu_n,\sigma^2)$

$=\prod_{i=1}^{n}\prod_{j=1}^{k}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x_{ij}-\mu_i)^2\right)$

$\therefore L(\mu_1\sim\mu_n,\sigma^2\mid x_{ij}\text{s})=\prod_{i=1}^{n}\prod_{j=1}^{k}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x_{ij}-\mu_i)^2\right)$

$\log L(\mu_1\sim\mu_n,\sigma^2\mid x_{ij}\text{s})=\sum_{i=1}^{n}\sum_{j=1}^{k}\left\{-\frac{1}{2}\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}(x_{ij}-\mu_i)^2\right\}$

(a) $\frac{\partial}{\partial\sigma^2}\log L=\sum_{i=1}^{n}\sum_{j=1}^{k}\left(\frac{1}{2\sigma^4}(x_{ij}-\mu_i)^2-\frac{1}{2\sigma^2}\right)=0$

$\therefore\sum_{i=1}^{n}\sum_{j=1}^{k}(x_{ij}-\mu_i)^2=nk\sigma^2\quad\therefore\hat{\sigma}^2_{MLE}=\frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k}(x_{ij}-\hat{\mu}_i)^2$

$\frac{\partial}{\partial\mu_1}\log L=\sum_{j=1}^{k}\frac{1}{\sigma^2}(x_{1j}-\mu_1)=0\Rightarrow\quad\hat{\mu}_{1MLE}=\bar{x}_{1\cdot}=\frac{x_{11}+\dots x_{1k}}{k}$

(同樣道理 $\hat{\mu}_{iMLE}=\bar{x}_{i\cdot}=\frac{x_{i1}+x_{i2}+\dots x_{ik}}{k}$

$\therefore\hat{\sigma}^2_{MLE}=\frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k}(x_{ij}-\bar{x}_{i\cdot})^2$; $\hat{\mu}_{iMLE}=\bar{x}_{i\cdot}=\frac{x_{i1}+\dots x_{ik}}{k}$

(2) 證明 $\hat{\sigma}^2_{MLE}$ 並非 σ^2 之一致估計量

我們注意 x_{ij}s 為獨立，而且 $\sum_{j=1}^{k}\frac{(x_{ij}-\bar{x}_{i\cdot})^2}{\sigma^2}\sim\chi^2_{k-1}$.

$\therefore \sum_{i=1}^{n}\sum_{j=1}^{k}\frac{(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2} \sim \chi^2_{nk-n}$

$\therefore E\left[\sum_{i=1}^{n}\sum_{j=1}^{k}\frac{(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2}\right] = nk-n = n(k-1)$

$\therefore E\left[\sum_{i=1}^{n}\sum_{j=1}^{k}(X_{ij}-\bar{X}_{i\cdot})^2\right] = n(k-1)\sigma^2$

$\therefore E\left[\hat{\sigma}^2_{MLE}\right] = \frac{n(k-1)}{nk}\sigma^2 = \frac{k-1}{k}\sigma^2$

$\therefore V\left[\sum_{i=1}^{n}\sum_{j=1}^{k}\frac{(X_{ij}-\bar{X}_{i\cdot})^2}{\sigma^2}\right] = 2n(k-1)$

$\therefore V\left[\sum_{i=1}^{n}\sum_{j=1}^{k}(X_{ij}-\bar{X}_{i\cdot})^2\right] = 2n(k-1)\sigma^4$

$V\left[\underbrace{\frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k}(X_{ij}-\bar{X}_{i\cdot})^2}_{\hat{\sigma}^2_{MLE}}\right] = \frac{2(k-1)}{nk^2}\sigma^4$

$\therefore \lim_{n\to\infty} V[\hat{\sigma}^2_{MLE}] = 0 \quad \therefore \hat{\sigma}^2_{MLE} \xrightarrow{p} \frac{k-1}{k}\sigma^2$ (∵ 柴比雪夫)

在此 k 是個固定的常數（跟 n 無關）

$\therefore \hat{\sigma}^2_{MLE}$ 為 $\frac{k-1}{k}\sigma^2$ 之一致估計量，而並非 σ^2 之一致估計量

(3) 根據 (2) 的討論 $\frac{k}{k-1}(\hat{\sigma}^2_{MLE}) \xrightarrow{p} \sigma^2$

(∵ $X_n \xrightarrow{p} c_0 \quad aX_n \xrightarrow{p} ac_0$)

$\therefore C = \frac{k}{k-1}$

(23)

注 $V_p(r) = \dfrac{\pi^{\frac{p}{2}}}{\Gamma\left(\frac{p}{2}+1\right)} r^p$

$\boxed{10}$ $V_p(r) \overset{\text{def}}{=}$ p維空間之球的體積(半徑=r)

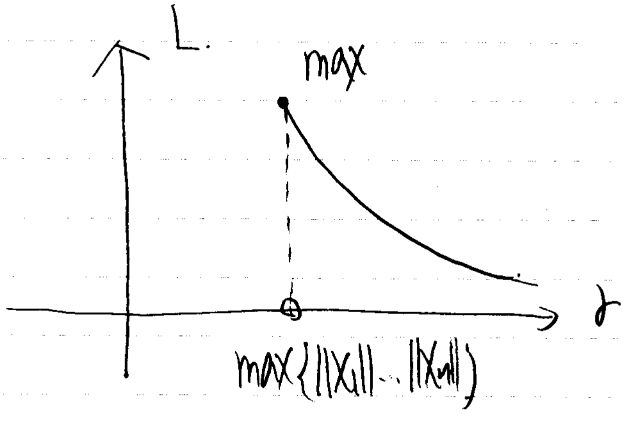
$$f(x|r) = \frac{1}{V_p(r)} \cdot I_{[0,r]}(\|x\|)$$

$$\therefore f(x_1 \cdots x_n | r) = \left\{\frac{1}{V_p(r)}\right\}^n \cdot I_{[0,r]}(\|x_1\|) \cdot I_{[0,r]}(\|x_2\|) \cdots I_{[0,r]}(\|x_n\|)$$

$$= \left\{\frac{1}{V_p(r)}\right\}^n \cdot I_{[0,r]}(\max\{\|x_1\|, \cdots \|x_n\|\})$$

$$L(r|x_1 \sim x_n) \overset{\text{def}}{=} \left\{\frac{1}{V_p(r)}\right\}^n \cdot I_{[\max\{\|x_1\| \cdots \|x_n\|\}, \infty)}(r)$$

$\therefore$



$$\therefore \hat{r}_{MLE} = \max\{\|x_1\| \cdots \|x_n\|\}$$

注 $\|x\| = \left(\sum_{j=1}^{p} x_j^2\right)^{\frac{1}{2}}$ $\quad \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$

接著考慮 $\hat{r}_{MLE}$ 之極限分布

$$\Pr(\hat{r}_{MLE} \leq t) = \Pr(\|X_1\| \leq t, \|X_2\| \leq t \cdots, \|X_n\| \leq t)$$

$$\underset{(\text{cdf})}{F_{\hat{r}}(t)} = \begin{cases} \left\{\dfrac{V_p(t)}{V_p(r)}\right\}^n = \left(\dfrac{t}{r}\right)^{np} & (0 \leq t \leq r) \\ 1 & (t > r) \end{cases}$$

$$\lim_{n\to\infty} \Pr\left(\hat{r}_{MLE} \leq r - \frac{t}{n}\right) = \lim_{n\to\infty}\left(1 - \frac{t}{rn}\right)^{np} = \lim_{n\to\infty}\left\{\left(1 - \frac{t}{nr}\right)^{-\frac{nr}{t}}\right\}^{-\frac{pt}{r}}$$

$$= e^{-\frac{p}{r}t}$$

$$\therefore \lim_{n\to\infty} \Pr\left(\hat{r}_{MLE} > r - \frac{t}{n}\right) = \lim_{n\to\infty} \Pr\left(\frac{t}{n} > r - \hat{r}_{MLE}\right) = 1 - e^{-\frac{p}{r}t}$$

$$\Leftrightarrow \lim_{n\to\infty} \Pr\left(n(r - \hat{r}_{MLE}) < t\right) = 1 - e^{-\frac{p}{r}t}$$

由此可知 $n(r - \hat{r}_{MLE}) \xrightarrow{d} \exp\left(\frac{p}{r}\right)$

(25)

11 為了避免混淆, $X_{m+1} \sim X_{m+n}$ 寫成 $Y_1 \sim Y_n$

$\therefore X_1 \sim X_m \overset{iid}{\sim} N(\mu_1, \sigma^2)$. $Y_1 \sim Y_n \overset{iid}{\sim} N(\mu_2, \sigma^2)$

$(\mu_1 \le \mu_2)$

(a) $f(x_1 \sim x_m, y_1 \sim y_n \mid \mu_1, \mu_2, \sigma^2)$

$$= \left\{ \prod_{j=1}^{m} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \exp\left(\frac{-1}{2\sigma^2}(x_j-\mu_1)^2 \right) \right\} \cdot \left\{ \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-1}{2\sigma^2}(y_j-\mu_2)^2 \right) \right\}$$

$$\log f(x_1 \sim x_m, y_1 \sim y_n \mid \mu_1, \mu_2, \sigma^2) = \frac{-1}{2}(m+n)\log(2\pi\sigma^2) - \sum_{j=1}^{m} \frac{1}{2\sigma^2}(x_j-\mu_1)^2 - \sum_{j=1}^{n} \frac{1}{2\sigma^2}(y_j-\mu_2)^2$$

$$L(\mu_1, \mu_2, \sigma^2) \overset{def}{=} \frac{-1}{2}(m+n)\log(2\pi\sigma^2) - \sum_{j=1}^{m} \frac{1}{2\sigma^2}(x_j-\mu_1)^2 - \sum_{j=1}^{n} \frac{1}{2\sigma^2}(y_j-\mu_2)^2$$

(這題我們只對 μ_1 和 μ_2 的 MLE 有興趣, 而且 μ_1, μ_2 的 MLE 跟 σ^2 無關, 所以以下將 σ^2 當一個固定常數)

$\frac{\partial L}{\partial \mu_1} = \sum_{j=1}^{m} \frac{(x_j-\mu_1)}{\sigma^2} = 0 \Rightarrow \mu_1 = \bar{x}$

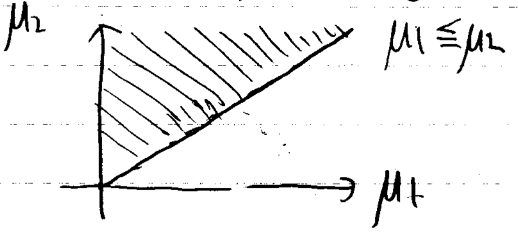
$\frac{\partial^2 L}{\partial \mu_1^2} = \frac{-n}{\sigma^2}$

同樣道理, $\frac{\partial L}{\partial \mu_2} = 0 \Rightarrow \mu_2 = \bar{y}$ $\frac{\partial^2 L}{\partial \mu_2^2} = \frac{-m}{\sigma^2}$

另外 $\frac{\partial^2 L}{\partial \mu_1 \partial \mu_2} = 0$

No.
Date

$\therefore$ Hessian $\begin{bmatrix} \frac{-n}{\sigma^2} & 0 \\ 0 & \frac{-m}{\sigma^2} \end{bmatrix}$ … 顯然為負定矩陣



⇓

接下來考慮L的最大值。

• case I… 若 $\overline{X} \leq \overline{Y}$ … $\begin{cases} \hat{\mu}_{1\,MLE} = \overline{X} \\ \hat{\mu}_{2\,MLE} = \overline{Y} \end{cases}$ 使得L最大

• case II… 若 $\overline{X} > \overline{Y}$ … 最大值發生於 $\mu_1 = \mu_2$ 上。

⇒ 我們應利用 Lagrange 未定乘數法

⇒ 但在此直接 $\mu = \mu_1 = \mu_2$ 比較快。

$$L(\mu) \overset{\text{re-def}}{=} \frac{-1}{2}(m+n)\log(2\pi\sigma^2) - \sum_{j=1}^{m}\frac{1}{2\sigma^2}(x_j-\mu)^2 - \sum_{j=1}^{n}\frac{1}{2\sigma^2}(y_j-\mu)^2$$

$$\frac{\partial L}{\partial \mu} = \sum_{j=1}^{m}\frac{1}{\sigma^2}(x_j-\mu) + \sum_{j=1}^{n}\frac{1}{\sigma^2}(y_j-\mu) = 0$$

$$\Leftrightarrow (m+n)\mu = m\overline{X} + n\overline{Y} \qquad \therefore \mu = \frac{m\overline{X}+n\overline{Y}}{m+n}$$

綜合 case I, II … $\begin{cases} \text{if } \overline{X} \leq \overline{Y} \Rightarrow \hat{\mu}_{1\,MLE} = \overline{X} \quad \hat{\mu}_{2\,MLE} = \overline{Y} \\ \text{if } \overline{X} > \overline{Y} \Rightarrow \hat{\mu}_1 = \hat{\mu}_2 = \frac{m\overline{X}+n\overline{Y}}{m+n} \end{cases}$

27

(PS) 老師後來把「asymtotic distribution」改為「distribution」
但 $\boxed{\text{case I}}$ 應該考慮 $m\to\infty$, $n\to\infty$ 的情況.
故此, 依照原本的題目作答. ($\boxed{\text{case II}}$ 應該不用考慮 asymtotic)

11 - part 2

(b) 接下來考慮 $(\hat{\mu}_1, \hat{\mu}_2)$ 之極限分佈 … (考慮聯合分佈)

$$Pr(\hat{\mu}_1 \leq a, \hat{\mu}_2 \leq b) = {}^{①}Pr(\hat{\mu}_1 \leq a, \hat{\mu}_2 \leq b, \bar{X} \leq \bar{Y}) + {}^{②}Pr(\hat{\mu}_1 \leq a, \hat{\mu}_2 \leq b, \bar{X} > \bar{Y})$$

$$= \overset{①}{Pr(\bar{X} \leq a, \bar{Y} \leq b, \bar{X} \leq \bar{Y})} + \overset{②}{Pr\left(\frac{m\bar{X}+n\bar{Y}}{m+n} \leq \min\{a,b\}, \bar{X} > \bar{Y}\right)}$$

$\boxed{\text{case I}}$ $\mu_1 < \mu_2$ …

$$① = Pr(\bar{X} \leq a, \bar{Y} \leq b, \bar{X} \leq \bar{Y}) = Pr(\bar{X} \leq a, \bar{Y} \leq b) - Pr(\bar{X} \leq a, \bar{Y} \geq b, \bar{X} > \bar{Y})$$

$$\geq Pr(\bar{X} \leq a, \bar{Y} \leq b) - Pr(\bar{X} > \bar{Y})$$

「$Pr(\bar{X} \leq a, \bar{Y} \leq b) - Pr(\bar{X} > \bar{Y}) \leq ① \leq Pr(\bar{X} \leq a, \bar{Y} \leq b)$」

$$Pr(\bar{X} > \bar{Y}) = Pr(W > 0) \quad \text{where } W = \bar{X} - \bar{Y}$$

$$W \sim N\left(\mu_1 - \mu_2, \left(\frac{1}{m} + \frac{1}{n}\right)\sigma^2\right)$$

且, $m\to\infty$, $n\to\infty$ 時, $\left(\frac{1}{m}+\frac{1}{n}\right)\sigma^2 \to 0$

$\therefore W \xrightarrow{P} \mu_1 - \mu_2 < 0$

$\therefore Pr(\bar{X} > \bar{Y}) \to 0$

$\therefore m\to\infty, n\to\infty \Rightarrow ① \to Pr(\bar{X} \leq a, \bar{Y} \leq b)$

② …… $② \leq Pr(\bar{X} > \bar{Y})$ 同樣道理, $Pr(\bar{X} > \bar{Y}) \to 0$

$\therefore ② \to 0$

$$\bar{X} \sim N\left(\mu_1, \frac{\sigma^2}{m}\right) \quad \bar{Y} \sim N\left(\mu_2, \frac{\sigma^2}{n}\right)$$

$$\left(\frac{m\bar{X}+n\bar{Y}}{m+n} \sim N\left(\mu_1, \frac{\sigma^2}{m+n}\right) \text{ if } \mu_1 = \mu_2 \right)$$

∴ $\mu_1 < \mu_2$ 時, $m, n \to \infty$ $\Pr(\hat{\mu}_1 \le a, \hat{\mu}_2 \le b) \leftrightarrow \Pr(\bar{X} \le a, \bar{Y} \le b)$

$$\Pr(\hat{\mu}_1 \le a) \xrightarrow{m,n\to\infty} \Pr(\bar{X} \le a), \quad \Pr(\hat{\mu}_2 \le b) \xrightarrow{m,n\to\infty} \Pr(\bar{Y} \le b)$$

故 $\hat{\mu}_1$ 與 $\hat{\mu}_2$ 之極限分佈與 $\bar{X}$ 與 $\bar{Y}$ 之分佈一致,

$$\therefore \hat{\mu}_1 \xrightarrow{d} N\left(\mu_1, \frac{\sigma^2}{m}\right) \quad \hat{\mu}_2 \xrightarrow{d} N\left(\mu_2, \frac{\sigma^2}{n}\right)$$

$\boxed{\text{Case II}}$ $\mu_1 = \mu_2$ … $\begin{cases} \text{if } \bar{X} \le \bar{Y} \ \cdots \ \hat{\mu}_{1MLE} = \bar{X}, \ \hat{\mu}_{2MLE} = \bar{Y} \\ \text{if } \bar{X} > \bar{Y} \ \cdots \ \hat{\mu}_{1MLE} = \hat{\mu}_{2MLE} = \dfrac{m\bar{X}+n\bar{Y}}{m+n} \end{cases}$

我們應該先考慮 $\Pr(\bar{X} \le \bar{Y})$ 與 $\Pr(\bar{X} > \bar{Y})$ 之機率.

$$W \overset{\text{def}}{=} \bar{X} - \bar{Y} \quad W \sim N\left(0, \left(\frac{1}{m}+\frac{1}{n}\right)\sigma^2\right) \quad (\because \mu_1 = \mu_2)$$

$$\Pr(W<0) = \frac{1}{2}, \quad \Pr(W \ge 0) = \frac{1}{2} \quad (\forall m, n \in \mathbb{Z})$$

$$\therefore \hat{\mu}_{1MLE} = \begin{cases} \bar{X} & \left(\frac{1}{2}\right) \\ \dfrac{m\bar{X}+n\bar{Y}}{m+n} & \left(\frac{1}{2}\right) \end{cases} \quad \hat{\mu}_{2MLE} = \begin{cases} \bar{Y} & \left(\frac{1}{2}\right) \\ \dfrac{m\bar{X}+n\bar{Y}}{m+n} & \left(\frac{1}{2}\right) \end{cases}$$

由此可知, $\hat{\mu}_{1MLE}$ 和 $\hat{\mu}_{2MLE}$ 為 Mixture Normal Distribution.

(但跟 case I 不同的是, $\Pr(\bar{X} \le \bar{Y}) = \Pr(\bar{X} > \bar{Y}) = \frac{1}{2}$ for all $m, n \in \mathbb{Z}$
所以不用考慮 asymptotic Distribution.)

$$\begin{cases} \cdot \ \hat{\mu}_{1MLE} \sim \overset{\frac{1}{2}}{N}\left(\mu_1, \frac{\sigma^2}{m}\right) \ \& \ \overset{\frac{1}{2}}{N}\left(\mu_1, \frac{\sigma^2}{m+n}\right) \ \left(\overset{p_1=}{\frac{1}{2}}, \overset{p_2=}{\frac{1}{2}}\right) \text{ 之混合常態分布} \\ \cdot \ \hat{\mu}_{2MLE} \sim \overset{\frac{1}{2}}{N}\left(\mu_1, \frac{\sigma^2}{n}\right) \ \& \ \overset{\frac{1}{2}}{N}\left(\mu_1, \frac{\sigma^2}{m+n}\right) \ \left(\frac{1}{2}, \frac{1}{2}\right) \text{ 之混合常態分布} \end{cases}$$

$\boxed{12}$ 先求 $E[X_1-t \mid X_1>t]$

考慮 $X_1-t \mid X_1>t$ 之分佈

$$Pr(X_1-t<x \mid X_1>t) = \frac{Pr(t<X_1<t+x)}{Pr(X_1>t)}$$

$$= \frac{\int_t^{t+x} \lambda e^{-\lambda x}\,dx}{\int_t^{\infty} \lambda e^{-\lambda x}\,dx} = \frac{\lambda(e^{-\lambda t}-e^{-\lambda t-\lambda x})}{\lambda e^{-\lambda t}}$$

$$= (1-e^{-\lambda x})$$

$$\therefore \frac{d}{dx} Pr(X_1-t<x \mid X_1>t) = \lambda e^{-\lambda x}$$

由此可知 $X_1-t \mid X_1>t \sim \exp(\lambda)$

$$\therefore E[X_1-t \mid X_1>t] = \frac{1}{\lambda}$$

$\therefore$ 求 $\frac{1}{\lambda}$ 之 MLE. 根據 MLE 之不變性 (invariant property)

$\frac{1}{\lambda}$ 之 MLE $= \frac{1}{\hat{\lambda}}$

$X_1 \sim X_n$ 之聯合密度函數 $f(x_1 \sim x_n \mid \lambda) = \lambda^n \exp(-\lambda(x_1+\cdots+x_n))$

$(x_1 \sim x_n \geq 0)$

$L(\lambda) = \lambda^n \exp(-\lambda(x_1+\cdots+x_n))$ $\quad \frac{\partial}{\partial \lambda} \log L(\lambda) = \frac{n}{\lambda} - (x_1+\cdots+x_n)$

No

Date

$$\frac{\partial \log L}{\partial \lambda}=0 \Rightarrow \lambda=\frac{1}{(\overline{X})}$$

λ	0		$\frac{1}{\overline{X}}$	
$\frac{\partial \log L}{\partial \lambda}$		$+$	0	$-$
L		$\nearrow$	max	$\searrow$

$\therefore \hat{\lambda}_{MLE}=\frac{1}{\overline{X}}$

$$\therefore \frac{1}{\lambda}\text{ 之 } MLE=\frac{1}{\hat{\lambda}}=\overline{X}=\frac{X_1+X_2+\cdots X_n}{n}$$

$E[X_1]=\frac{1}{\lambda}<\infty$ $\therefore$ 根據 Khinchine 之弱大數法則

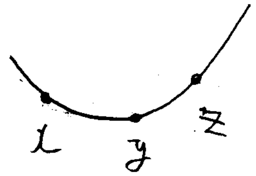
$$\frac{X_1+X_2+\cdots X_n}{n}\xrightarrow{P}\frac{1}{\lambda}\quad(\because X_1\sim X_n \text{ iid},\ |E[X_i]|<\infty)$$

$$\therefore \frac{1}{\hat{\lambda}}=\frac{X_1+\cdots X_n}{n}\xrightarrow{P}\frac{1}{\lambda}\quad(\text{退化})$$

或者 $V[X_i]=\frac{1}{\lambda^2}<\infty$ $\therefore$ 根據中央極限定理,

$$\sqrt{n}\left(\frac{X_1+\cdots X_n}{n}-\frac{1}{\lambda}\right)\xrightarrow{d}N\left(0,\frac{1}{\lambda^2}\right)$$

$$\therefore \sqrt{n}\left(\frac{1}{\hat{\lambda}}-\frac{1}{\lambda}\right)\xrightarrow{d}N\left(0,\frac{1}{\lambda^2}\right)$$



$\boxed{\text{NOTE}}$ $\boxed{2}$ $x \leqq z$

① $f((1-t)x+tz) \leqq (1-t)f(x) + tf(z)$

$\Leftrightarrow f((1-t)x+tz) - f(x) \leqq t(f(z)-f(x))$

$\Leftrightarrow \dfrac{f((1-t)x+tz) - f(x)}{t(z-x)} \leqq \dfrac{f(z)-f(x)}{z-x}$

$y \overset{\text{put}}{=} x + t(z-x)$

$\Leftrightarrow \dfrac{f(y)-f(x)}{y-x} \leqq \dfrac{f(z)-f(x)}{z-x} \quad (x \leqq y \leqq z)$

② $f((1-t)x+tz) \leqq (1-t)f(x) + tf(z)$

$\Leftrightarrow f((1-t)x+tz) - f(z) \leqq (1-t)(f(x)-f(z))$

$\Leftrightarrow f((1-t)x+tz) - f(z) \geqq (1-t)(f(z)-f(x))$

$\Leftrightarrow \dfrac{f((1-t)x+tz) - f(z)}{(1-t)(z-x)} \geqq \dfrac{f(z)-f(x)}{z-x}$

$y \overset{\text{put}}{=} (1-t)x+tz$

$\Leftrightarrow \dfrac{f(z)-f(y)}{z-y} \geqq \dfrac{f(z)-f(x)}{z-x} \quad (x \leqq y \leqq z)$

綜合①,②：$x \leq y \leq z$ ，$\dfrac{f(y)-f(x)}{y-x} \leq \dfrac{f(z)-f(x)}{z-x} \leq \dfrac{f(z)-f(y)}{z-y}$

高等統計推論II 第4次作業 R05246013 森元俊成

1

求MLE（θ使得$L(\theta)$最大，$\theta\in[0,1]$）

$X=-2 \cdots L(\theta)=\frac{1-\theta}{4}$　　$\underset{\theta}{\arg\max}\,L(\theta)=0$

$X=-1 \cdots L(\theta)=\frac{\theta}{12}$　　$\underset{\theta}{\arg\max}\,L(\theta)=1$

$X=0 \cdots L(\theta)=\frac{1}{2}$　　$\underset{\theta}{\arg\max}\,L(\theta)=\forall c\in[0,1]$

$X=1 \cdots L(\theta)=\frac{3-\theta}{12}$　　$\underset{\theta}{\arg\max}\,L(\theta)=0$

$X=2 \cdots L(\theta)=\frac{\theta}{4}$　　$\underset{\theta}{\arg\max}\,L(\theta)=1$

$$\hat{\theta}_{MLE}=\begin{cases}0 & (X=-2)\\ 1 & (X=-1)\\ \forall c\in[0,1] & (X=0)\\ 0 & (X=1)\\ 1 & (X=2)\end{cases}$$

$\cdots$ $\hat{\theta}_{MLE}$並不唯一.

求θ的不偏估計量

$$E[X]=\frac{1-\theta}{4}\cdot(-2)+\frac{\theta}{12}(-1)+\frac{1}{2}\cdot(0)+\frac{3-\theta}{12}\cdot 1+\frac{\theta}{4}(2)$$

$$=\frac{\theta-1}{2}-\frac{\theta}{12}+\frac{3-\theta}{12}+\frac{\theta}{2}$$

$$=\frac{6\theta-6-\theta+3-\theta+6\theta}{12}$$

$$=\frac{10\theta-3}{12}$$

$\therefore$ $E\left[\frac{6X}{5}+\frac{3}{10}\right]=\theta$

$\therefore \frac{6}{5}X + \frac{3}{10}$ 為 θ 的不偏估計量,

$$E[\hat{\theta}_{MLE}] = \frac{1-\theta}{4}(0) + \frac{\theta}{12}(1) + \frac{1}{2}(c) + \frac{3-\theta}{12}(0) + \frac{\theta}{4}(1)$$

$$= \frac{\theta}{12} + \frac{c}{2} + \frac{\theta}{4} = \frac{\theta}{3} + \frac{c}{2}$$

c 為跟 θ 無關的常數, $\forall c$ 無法使得

$E[\hat{\theta}_{MLE}] = \theta$. $\hat{\theta}_{MLE}$ 不可能 θ 之不偏估計量.

Date Page

$\boxed{2}$ $X_1, X_2, \cdots X_n \sim \text{bernoulli}(p)$

(a) $Pr(X_1 = x_1, \cdots, X_n = x_n \mid p)$

$= p^{(x_1+x_2+\cdots+x_n)} (1-p)^{(n-x_1-x_2\cdots-x_n)}$

$l(p) = \ln Pr(X_1 = x_1, \cdots, X_n = x_n)$

$= (x_1+x_2+\cdots+x_n)\ln p + (n-x_1-x_2\cdots-x_n)\ln(1-p)$

$l'(p) = \frac{1}{p}(x_1+x_2+\cdots+x_n) - \frac{1}{1-p}(n-x_1-x_2\cdots-x_n) = 0$

$\Leftrightarrow (1-p)(x_1+x_2+\cdots+x_n) - p(n-x_1-x_2\cdots-x_n) = 0$

$\Leftrightarrow p = \bar{X}$

$l''(p) = \frac{-1}{p^2}(x_1+\cdots+x_n) - \frac{1}{(1-p)^2}(n-x_1-x_2\cdots-x_n) < 0$

$\therefore \hat{p}_{MLE} = \bar{X}$

根據中央極限定理,

$\sqrt{n}\left(\frac{X_1+X_2+\cdots+X_n}{n} - p\right) \xrightarrow{d} N(0, p(1-p))$

$(E[X_i] = p \quad V[X_i] = p(1-p))$

$\hat{\theta}_{MLE} = \hat{p}_{MLE}(1-\hat{p}_{MLE})$ （∵ MLE不變性 ⇒ 第5次作業 (1)）

$g(p) = p(1-p)$ $\quad g'(p) = 1-2p$

利用 δ-method

$\sqrt{n}(g(\hat{p}) - g(p)) \xrightarrow{d} N(0, p(1-p)\cdot\{g'(p)\}^2)$

$\therefore \sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N(0, p(1-p)(1-2p)^2)$ $(p \neq \frac{1}{2})$

(b) 自己重新定義：$h(p)$ $(h(\frac{1}{2}) = g(\frac{1}{2}) = \frac{1}{4})$

例如：取 $h(x) = x^2$. $\quad h'(x) = 2x$

$\therefore \sqrt{n}(\underset{\hat{\theta}_{2\,MLE}}{\underset{\parallel}{h(\hat{p})}} - \underset{\theta_2}{\underset{\parallel}{h(p)}}) \xrightarrow{d} N(0, p(1-p)\cdot(h'(p))^2)$

$\qquad\qquad = N(0, p(1-p)\cdot 4)$

$p = \frac{1}{2}$ ∴ $\sqrt{n}(\hat{\theta}_{2\,MLE} - \theta_2) \xrightarrow{d} N(0, 1)$

Date Page

$\boxed{3}$ $F(x,y) = x \cdot y \cdot (\min\{x,y\})^{\theta}$

- $x > y$ … $F(x,y) = x \cdot y^{\theta+1}$
- $x < y$ … $F(x,y) = x^{\theta+1} \cdot y$

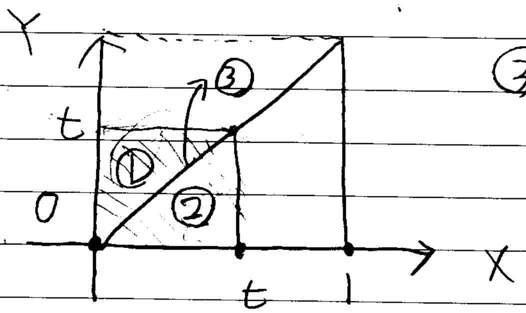
$$\frac{\partial^2 F}{\partial x \partial y} = \begin{cases} (\theta+1)\, y^{\theta} & (x>y) \\ (\theta+1)\, x^{\theta} & (x<y) \end{cases}$$
$(X \neq Y)$

$$= (\theta+1) \cdot (\min\{x,y\})^{\theta}$$

這個分佈其實是兩種連續分佈的混合分佈

($X=Y$ 的部分與 $X \neq Y$ 的部分)

接著求 $X=Y$ 的分佈：



③：$X=Y$ 的直線上

① 的機率 $(0<X<Y<t)$

$$= \int_{0<x<y<t} (\theta+1)\, x^{\theta}\, dx\, dy$$

$$= \int_{0<y<t} \left[x^{\theta+1}\right]_0^{y} dy$$

$$= \int_0^t y^{\theta+1} dy = \frac{t^{\theta+2}}{\theta+2}$$

② 的概率 $(0 < Y < X < t)$

同理, $= \dfrac{t^{\theta+2}}{\theta+2}$

①+②+③ 的概率 $(X, Y \leq t)$

$F(t, t) = t^{\theta+2}$

∴ ③ $(= Pr(X = Y \leq t))$ 的概率

$= (①+②+③) - (①+②)$

$$= t^{\theta+2} - \frac{2t^{\theta+2}}{\theta+2} = \frac{\theta}{\theta+2} t^{\theta+2}$$

另外, 代入 $t=1$ 得知 $\begin{cases} Pr(X = Y) = \dfrac{\theta}{\theta+2} \\ Pr(X \neq Y) = \dfrac{2}{\theta+2} \end{cases}$

- $X \neq Y$ 下的概率密度.. $\left(\dfrac{\theta+2}{2}\right) \cdot (\theta+1) (\min\{x, y\})^{\theta}$
- $X = Y$ 下的概率密度.. $Pr(X = Y \leq x \mid X = Y)$

$$= x^{\theta+2}$$

$$\because \frac{d}{dx}x^{\theta+2} = (\theta+2)x^{\theta+1} \quad \therefore \text{Beta}(\theta+2, 1)$$

總而言之，

- $(X,Y) \mid X \neq Y \sim f_{X\neq Y}(x,y) = \left(\frac{\theta+2}{2}\right)(\theta+1)\left(\min\{x,y\}\right)^{\theta}$

  $\Pr(X \neq Y) = \frac{2}{\theta+2}$

- $(X,Y) \mid X = Y \sim f_{X=Y}(x) = (\theta+2)x^{\theta+1}$

  $\Pr(X = Y) = \frac{\theta}{\theta+2}$

（但 $0<x<1$, $0<y<1$）

(a) 根據以上的討論，完整的機率密度函數為

$$f(x,y) = f_{X\neq Y}(x,y)\cdot \Pr(X\neq Y) + f_{X=Y}(x)\cdot \Pr(X=Y)$$

$$= (\theta+1)\left(\min\{x,y\}\right)^{\theta} I(X\neq Y) + \theta x^{\theta+1} I(X=Y)$$

$$= \left\{(\theta+1)\min\{x,y\}^{\theta}\right\}^{1-I(X=Y)} \cdot \left\{\theta x^{(\theta+1)}\right\}^{I(X=Y)}$$

故概似比函數為

$$L(\theta) = \prod_{j=1}^{n} \left\{(\theta+1)\min\{x_j,y_j\}^{\theta}\right\}^{1-I(X_j=Y_j)} \cdot \left(\theta x_j^{(\theta+1)}\right)^{I(X_j=Y_j)}$$

(b) 接著考慮 $\hat{\theta}_{MLE}$ (但應該不漂亮)

為了簡單, $I_j \overset{def}{=} I(X_j = Y_j)$

$$\ell(\theta) = \ln L(\theta) = \sum_{j=1}^{n} \{ (1-I_j)\ln(\theta+1) + (1-I_j)\cdot\theta\cdot\ln\min\{x_j, y_j\} + I_j \ln\theta + I_j\cdot(\theta+1)\ln x_j \}$$

$$\ell'(\theta) = \sum_{j=1}^{n} \left(\frac{1-I_j}{\theta+1} + (1-I_j)\ln\min\{x_j, y_j\} + \frac{I_j}{\theta} + I_j \ln x_j \right)$$

$$\ell''(\theta) = \frac{-1}{(\theta+1)^2}\sum_{j=1}^{n}(1-I_j) - \frac{1}{\theta^2}\sum_{j=1}^{n} I_j < 0$$

$$\ell'(\theta) = 0 \Rightarrow \theta(\theta+1)\ell'(\theta) = 0$$

$$\Rightarrow \sum_{j=1}^{n} \{ \underbrace{\theta(1-I_j) + (\theta+1)I_j}_{\to\ \theta + I_j} + \theta(\theta+1)\{(1-I_j)\ln\min\{X_j, Y_j\} + I_j \ln X_j\} \}$$

$$= \sum_{j=1}^{n} \{ \theta^2 \left((1-I_j)\ln\min\{X_j, Y_j\} + I_j \ln X_j \right) + \theta\left((1-I_j)\ln\min\{X_j, Y_j\} + I_j \ln X_j + 1 \right) + I_j \} = 0 \Leftrightarrow a\theta^2 + b\theta + c = 0$$

where
$$\begin{cases} a = \sum_{j=1}^{n} \{ (1-I_j)\ln\min\{X_j, Y_j\} + I_j \ln X_j \} \\ b = \sum_{j=1}^{n} \{ (1-I_j)\ln\min\{X_j, Y_j\} + I_j \ln X_j + 1 \} \\ c = \sum_{j=1}^{n} I_j \end{cases}$$

∴ $\hat{\theta}_{MLE}$ 為 $a\theta^2+b\theta+c=0$ 之根.

$$\begin{cases} \cdot\ a=\sum_{j=1}^{n}(1-I_j)\ln\min\{X_j,Y_j\}+\sum_{j=1}^{n}I_j\ln X_j \\ \cdot\ b=n+\sum_{j=1}^{n}(1-I_j)\ln\min\{X_j,Y_j\}+\sum_{j=1}^{n}I_j\ln X_j \\ \cdot\ c=\sum_{j=1}^{n}I_j \qquad (I_j=I(X_j=Y_j)) \end{cases}$$

(c) $\hat{\theta}_{MLE}$ 的漸近分佈

我們假設 Regularity Condition 成立:

$$l''(\theta)=\frac{-1}{(\theta+1)^2}\sum_{j=1}^{n}(1-I_j)-\frac{1}{\theta^2}\sum_{j=1}^{n}I_j$$

$$E[I_j]=Pr(I_j=1)=Pr(X_j=Y_j)=\frac{\theta}{\theta+2}$$

$$-E[l''(\theta)]=E\left[\frac{1}{(\theta+1)^2}\sum_{j=1}^{n}(1-I_j)+\frac{1}{\theta^2}\sum_{j=1}^{n}I_j\right]$$

$$=\left\{\frac{1}{(\theta+1)^2}\cdot\left(\frac{2}{\theta+2}\right)+\frac{1}{\theta^2}\cdot\frac{\theta}{\theta+2}\right\}\cdot n$$

$$=n\cdot\left(\frac{2}{(\theta+1)^2(\theta+2)}+\frac{1}{\theta(\theta+2)}\right)$$

∴ $\hat{\theta}_{MLE}-\theta \xrightarrow{d} N(0,\ -E[l''(\theta)]^{-1})$

故 $\sqrt{n}(\hat{\theta}_{MLE}-\theta) \xrightarrow{d} N\left(0, \left\{\frac{2}{(\theta+1)^2(\theta+2)}+\frac{1}{\theta(\theta+2)}\right\}^{-1}\right)$

Date Page

$\boxed{4}$ $L(\theta|x) \overset{\text{def}}{=} \exp(\eta(\theta)Y(x) - B(\theta))h(x)$

$l(\theta) \overset{\text{def}}{=} \ln L(\theta|x) = \eta(\theta)Y(x) - B(\theta) + \ln h(x)$

- $\frac{\partial}{\partial\theta} l(\theta) = \eta'(\theta)Y(x) - B'(\theta) = 0$

$\therefore \hat{\theta}_{MLE}(x)$ 満足 $\eta'(\hat{\theta})Y(x) - B'(\hat{\theta}) = 0 \cdots (*)$

$l(\hat{\theta}) - l(\theta_0) = \eta(\hat{\theta})Y(x) - B(\hat{\theta}) + \ln h(x)$

$\quad - \eta(\theta_0)Y(x) + B(\theta_0) - \ln h(x)$

$= \eta(\hat{\theta})Y(x) - \eta(\theta_0)Y(x) - B(\hat{\theta}) + B(\theta_0)$

$\frac{d}{dY}(l(\hat{\theta}) - l(\theta_0)) = \frac{dx}{dY} \cdot \frac{d}{dx}(\eta(\hat{\theta})Y(x) - \eta(\theta_0)Y(x) - B(\hat{\theta}) + B(\theta_0))$

$= \frac{dx}{dY} \cdot \{\underbrace{\eta'(\hat{\theta})\hat{\theta}'(x)Y(x) + \eta(\hat{\theta})Y'(x) - \eta(\theta_0)Y'(x) - B'(\hat{\theta})\hat{\theta}'(x)}\}$

$= \hat{\theta}'(x)\{\eta'(\hat{\theta})Y(x) - B'(\hat{\theta})\} + Y'(x)(\eta(\hat{\theta}) - \eta(\theta_0))$

$= Y'(x)(\eta(\hat{\theta}) - \eta(\theta_0))$ ($\because (*)$ の條件)

$= \underbrace{\frac{dx}{dY} \cdot Y'(x)}_{=1}(\eta(\hat{\theta}) - \eta(\theta_0))$

$= \eta(\hat{\theta}) - \eta(\theta_0) \geq 0$ (when $\hat{\theta} \geq \theta_0$)

$$\therefore \frac{d}{dx}\left(\ell(\hat{\theta}(x)) - \ell(\theta_0)\right) = \eta(\hat{\theta}_{MLE}(x)) - \eta(\theta_0) \geq 0$$

when $\hat{\theta}_{MLE} \geq \theta_0$

($\because \eta$ 為遞增函數)

Date Page

$\boxed{5}$

(a) $L(\theta|x_1,\dots,x_n) = \prod_{j=1}^{n} \frac{\theta}{(1+x_j)^{\theta+1}}$

$\ell(\theta|x_1,\dots,x_n) = \ln L(\theta|x_1,x_2,\dots,x_n)$

$= \sum_{j=1}^{n} \left(\ln\theta - (\theta+1)\ln(1+x_j) \right)$

$= n\ln\theta - (\theta+1)\sum_{j=1}^{n}\ln(1+x_j)$

$\frac{\partial \ell(\theta|x_1,\dots,x_n)}{\partial\theta} = \frac{n}{\theta} - \sum_{j=1}^{n}\ln(1+x_j) = 0$

$\frac{\partial^2 \ell(\theta|x_1,\dots,x_n)}{\partial\theta^2} = \frac{-n}{\theta^2} < 0$

$\hat{\theta}_{MLE} = \frac{n}{\sum_{j=1}^{n}\ln(1+X_j)}$

(b) $Y \overset{def}{=} \ln(1+X)$

$\frac{dY}{dX} = \frac{1}{1+X} = \exp(-Y)$

$1 = \int_{x=0}^{x=\infty} \frac{\theta}{(1+x)^{\theta+1}}\,dx$

$= \int_{y=0}^{y=\infty} \theta \cdot e^{-(\theta+1)y} \cdot e^{y}\,dy$

$= \int_0^\infty \theta e^{-\theta y}\,dy \quad \therefore Y = \ln(1+X) \sim \exp(\theta) \quad (\text{mean } \frac{1}{\theta})$

根據中央極限定理,

$$\sqrt{n}\left(\frac{Y_1+Y_2+\cdots+Y_n}{n}-\frac{1}{\theta}\right)\xrightarrow{d}N\left(0,\frac{1}{\theta^2}\right)$$

利用 δ-method $g(x)=\frac{1}{x}\ (x>0)$ $g'(x)=\frac{1}{x^2}$

$$\sqrt{n}\left(g\left(\frac{Y_1+\cdots+Y_n}{n}\right)-g\left(\frac{1}{\theta}\right)\right)\xrightarrow{d}N\left(0,\frac{1}{\theta^2}g'\left(\frac{1}{\theta}\right)^2\right)$$

$$\Leftrightarrow \sqrt{n}\left(\hat{\theta}_{MLE}-\theta\right)\xrightarrow{d}N(0,\theta^2)$$

$\therefore$ $N(0,\theta^2)$

(c) $\sqrt{n}\left(g(\hat{\theta}_{MLE})-g(\theta)\right)\xrightarrow{d}N(0,\theta^2\cdot g'(\theta)^2)$

(δ-method)

$\therefore g'(\theta)=\pm\frac{1}{\theta}$

$\therefore g(\theta)=\pm\ln\theta+C$ (C: 常數)

Date Page

$\boxed{6}$ 利用 χ^2-test : H_0: ~Poisson ; H_1: not Poisson

k	0	1	2	3	4	5	6	7	8	9	10	11	...
O_k	275,	246,	178,	84,	67,	27,	17,	12,	1,	2,	1,	1	...

k	12	16	19
O_k	2,	1,	1

- $n = \sum_{k=0}^{19} O_k = 915.$
- $\hat{\lambda} = \sum_{k=0}^{19} k \cdot O_k / n = 1.6929$

$$E_k = n \cdot e^{-\hat{\lambda}} \cdot \frac{(\hat{\lambda})^k}{k!} \quad (k = 0 \sim 12)$$

$$E_{13} = n \cdot \left(1 - \sum_{k=0}^{12} e^{-\hat{\lambda}} \cdot \frac{(\hat{\lambda})^k}{k!}\right)$$

($k \geq 13$ 的部份合併在一起)

k	0	1	...	13
E_k	168.347	284.9942	...	$2.8811 \cdot 10^{-5}$

- 檢定統計量

$$T = \sum_{k=0}^{13} \frac{(E_i - O_i)^2}{E_i} = 1.605 \cdot 10^{5}$$

T的分佈 (H_0 T)

$T \sim \chi^2_{14-1-1} = \chi^2_{12}$ $\quad \chi^{2\,-1}_{12}(0.95) = 21.0261$

$T \gg 21.0261$

∴ 拒絕 H_0. (並不服從 Poisson)

Red Elephant

Date Page

$\boxed{7}$

(a) $E[e^{tX}] = \int_{x=-\infty}^{x=\infty} e^{tX} \cdot f(x)\,dx$

$= \int_{-\infty}^{\infty} e^{tX} \frac{\lambda}{2} \exp(-\lambda|X|)\,dx$

$= \int_{0}^{\infty} e^{tX} \frac{\lambda}{2} \exp(-\lambda X)\,dx + \int_{-\infty}^{0} e^{tX} \frac{\lambda}{2} \exp(\lambda X)\,dx$

$= \underbrace{\int_{0}^{\infty} \frac{\lambda}{2} \exp(-(\lambda-t)x)\,dx} + \underbrace{\int_{-\infty}^{0} \frac{\lambda}{2} \exp((\lambda+t)x)\,dx}$

$\frac{\lambda}{2} \cdot \frac{1}{\lambda-t} \cdot \underbrace{\int_{0}^{\infty} (\lambda-t) \exp(-(\lambda-t)x)\,dx}_{=1}$ $\qquad \int_{\infty}^{0} \frac{-\lambda}{2} \exp(-(\lambda+t)x)\,dx$

$= \int_{0}^{\infty} \frac{\lambda}{2} \exp(-(\lambda+t)x)\,dx$

$= \frac{\lambda}{2} \cdot \frac{1}{\lambda+t}$

$= \frac{\lambda}{2}\left(\frac{1}{\lambda-t} + \frac{1}{\lambda+t}\right) = \frac{\lambda}{2} \cdot \frac{2\lambda}{\lambda^2-t^2} = \underbrace{\frac{\lambda^2}{\lambda^2-t^2}} = M_X(t)$

(b) $U, V \overset{\text{i.i.d.}}{\sim} \exp(1)$

$E[e^{t(U-V)}] = E[e^{tU}] \cdot E[e^{(-t)V}]$

$E[e^{tU}] = \int_{u=0}^{u=\infty} e^{tu} \cdot e^{-u}\,du = \int_{u=0}^{u=\infty} e^{-(1-t)u}\,du$

$= \frac{1}{1-t}$

$$\therefore E\left[e^{tV}\right] = \frac{1}{1+t} \quad (\because \text{iid})$$

$$\therefore E\left[e^{t(U-V)}\right] = \frac{1}{1-t} \cdot \frac{1}{1+t} = \frac{1}{1-t^2} = \frac{1^2}{1-t^2} \quad (\lambda=1)$$

由(a)可知, $X|\lambda=1$ 與 $Y(=U-V)$之分佈相同.

(Levy 連續性定理)

$\therefore Y \sim \text{Double-Exp}(1) \ (= \text{Laplace}(1))$

(c) 若 $U, V \overset{iid}{\sim} \exp(\lambda)$ (mean $\frac{1}{\lambda}$)

$U-V \sim \text{Laplace}(\lambda)$ ($\text{Double-Exp}(\lambda)$)

$\therefore X_j \overset{def}{=} U_j - V_j$ ($\sim \text{Laplace}(\lambda)$)

$$E[X_j] = 0, \quad V[X_j] = V[U_j] + V[V_j] = \frac{1}{\lambda^2} + \frac{1}{\lambda^2} = \frac{2}{\lambda^2}$$

$$\begin{cases} \therefore \text{期望值} = 0 \\ \text{變異數} = \frac{2}{\lambda^2}, \quad \sigma^2 \overset{def}{=} \frac{2}{\lambda^2} \end{cases}$$

$$M_{W_n}(t) = E\left[e^{tW_n}\right] = E\left[\exp\left(t \frac{\sqrt{n}}{\sigma} \frac{1}{n}(X_1 + \cdots X_n)\right)\right]$$

$$= E\left[\exp\left(\frac{t}{\sigma\sqrt{n}}(X_1 + X_2 + \cdots X_n)\right)\right]$$

$$= E\left[\exp\left(\frac{t}{\sigma\sqrt{n}} X_1\right)\right]^n$$

Date Page

$$M_{W_n}(t) = E\left[\exp\left(\frac{t}{\sigma\sqrt{n}}X_1\right)\right]^n$$

$$= \left\{M_X\left(\frac{t}{\sigma\sqrt{n}}\right)\right\}^n$$

$$= \left\{\frac{\lambda^2}{\lambda^2 - \left(\frac{t}{\sigma\sqrt{n}}\right)^2}\right\}^n = \left\{\frac{\lambda^2}{\lambda^2 - \frac{t^2}{n\sigma^2}}\right\}^n \qquad \text{注} \left(\sigma^2 = \frac{2}{\lambda^2}\right)$$

$$= \left(\frac{\lambda^2}{\lambda^2 - \frac{\lambda^2 t^2}{2n}}\right)^n = \left(1 - \frac{t^2}{2n}\right)^n$$

$$\therefore M_{W_n}(t) = \left(1 - \frac{t^2}{2n}\right)^n$$

(d) $$\lim_{n\to\infty} M_{W_n}(t) = \lim_{n\to\infty}\left(1 - \frac{t^2}{2n}\right)^n$$

$$= \lim_{n\to\infty}\Big\{\underbrace{\left(1 - \frac{t^2}{2n}\right)^{-\frac{2n}{t^2}}}_{\to e}\Big\}^{-\frac{t^2}{2}}$$

$$= \exp\left(\frac{-t^2}{2}\right)$$ …… $N(0,1)$ 的 Moment Generating Function

根據 Levy 的連續性定理,

$$\lim_{n} M_{W_n}(t) = \exp\left(\frac{-t^2}{2}\right) \Rightarrow W_n \xrightarrow{d} N(0,1)$$

(五) $V[X_i] < \infty$

利用中央极限定理，$\sqrt{n}\left(\frac{X_1+\cdots X_n}{n}\right) \xrightarrow{d} N\left(0, \frac{2}{\lambda^2}\right)$

$\Rightarrow \sqrt{n}\left(\frac{\bar{X}}{\sqrt{V[X_i]}}\right) \xrightarrow{d} N(0,1)$ （得到相同的结论）

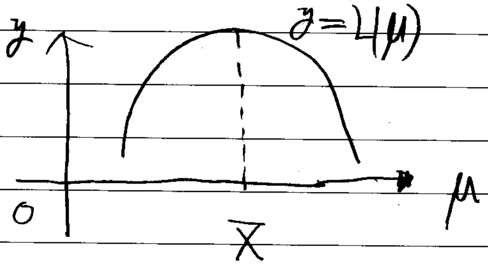
$\boxed{8}$ $X_1, X_2, \dots X_n \overset{iid}{\sim} N(\mu, 1)$

(a) $\prod_{j=1}^{n} f(x_j|\mu) = \prod_{j=1}^{n} \left(\frac{1}{\sqrt{2\pi}}\right) \cdot \exp\left(-\frac{1}{2}(x_j-\mu)^2\right)$

$\ell(\mu) = \ln \prod_{j=1}^{n} f(x_j|\mu) = \sum_{j=1}^{n} -\frac{1}{2}\ln(2\pi) - \frac{1}{2}(x_j-\mu)^2$

$\ell'(\mu) = \sum_{j=1}^{n} (x_j-\mu) = n(\bar{X}-\mu)$

$\ell''(\mu) = -n < 0$



① if $1 \leq \bar{X}$ $\hat{\mu}_{MLE} = \bar{X}$

② if $\bar{X} < 1$ $\hat{\mu}_{MLE} = 1$

$\therefore \hat{\mu}_{MLE} = \max\{\bar{X}, 1\}$

(b) 考慮 $\sqrt{n}(\hat{\mu}_{MLE}-\mu)$ 之 累積分布函數

$Pr(\sqrt{n}(\hat{\mu}_{MLE}-\mu) \leq \lambda)$

$= Pr(\hat{\mu}_{MLE} \leq (\mu + \frac{\lambda}{\sqrt{n}}))$

$= Pr(\bar{X}, 1 \leq \mu + \frac{\lambda}{\sqrt{n}})$

$$\Pr\left(\sqrt{n}(\bar{X}-\mu)\le x\right)\cdot I\left(1\le \mu+\frac{x}{\sqrt{n}}\right)$$

$$=\Phi(x)\cdot I_{[\sqrt{n}(1-\mu),\infty)}(x)$$

題目有錯，"$\mu>0$" 應為 "$\mu>1$"

$\mu>1$，$n\to\infty$ $\sqrt{n}(1-\mu)\to-\infty$

$$\lim_{n\to\infty}\Pr\left(\sqrt{n}(\hat{\mu}_{MLE}-\mu)\le x\right)=\lim_{n\to\infty}\Phi(x)I_{[\sqrt{n}(1-\mu),\infty)}(x)$$

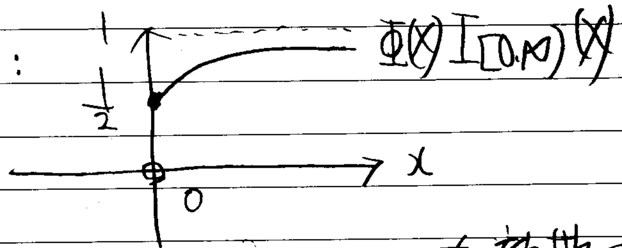
$$=\Phi(x)\,I_{(-\infty,\infty)}(x)=\Phi(x)$$

$\therefore \mu>1$ $\because$ $\sqrt{n}(\hat{\mu}_{MLE}-\mu)\xrightarrow{d}N(0,1)$

(c) 題目有錯 "$\mu=0$" 應為 "$\mu=1$"

$\Pr\left(\sqrt{n}(\hat{\mu}_{MLE}-\mu)\le x\right)=\Phi(x)\cdot I_{[0,\infty)}(x)$ (for all n)

cdf :



$\therefore$ 有離散的部分

$$\begin{cases}\Pr\left(\sqrt{n}(\hat{\mu}_{MLE}-\mu)=0\right)=\frac{1}{2}\\ \Pr\left(\sqrt{n}(\hat{\mu}_{MLE}-\mu)\le x\right)=\Phi(x)\quad(x>0)\end{cases}$$

(for all $n\ge1$)

Date Page

$\boxed{9}$

(a) $Pr(T \leq t) = \begin{cases} 0 \\ \left(\frac{t}{\theta}\right)^2 & (0<t<\theta) \\ 1 \end{cases}$

$$\frac{d}{dt} Pr(T \leq t) = \frac{2t}{\theta^2} \cdot I_{(0,\theta)}(t)$$

$$\therefore f_T(t) = \frac{2t}{\theta^2} I_{(0,\theta)}(t)$$

$$f_{X_1,X_2,t}(x_1,x_2,t) = \frac{1}{\theta^2} \cdot I_{(0,t)}(x_2) \cdot I_{\{t\}}(x_1)$$
$$+ \frac{1}{\theta^2} I_{(0,t)}(x_1) I_{\{t\}}(x_2)$$

$$\therefore f_{X_1,X_2|T}(X_1,X_2|T=t) = \frac{f_{x_1,x_2,t}(x_1,x_2,t)}{f_T(t)}$$

$$= \frac{\frac{1}{\theta^2}\left\{ I_{(0,t)}(x_2) I_{\{t\}}(x_1) + I_{(0,t)}(x_1) I_{\{t\}}(x_2) \right\}}{\frac{2t}{\theta^2} I_{(0,\theta)}(t)}$$

$(0<t<\theta$ を満足$)$

$$= \frac{1}{2t}\left\{ I_{(0,t)}(x_2) I_{\{t\}}(x_1) + I_{(0,t)}(x_1) I_{\{t\}}(x_2) \right\}$$

(b) 求 θ 之最小充份統計量

$(X_1,X_2) \neq (Y_1,Y_2)$ 之下

$\dfrac{f_{X_1,X_2}(X_1,X_2)}{f_{Y_1,Y_2}(Y_1,Y_2)}$ 與 θ 無關 $\Rightarrow \dfrac{\frac{1}{\theta^2}\cdot I_{(0,\theta)}(\max\{X_1,X_2\})}{\frac{1}{\theta^2} I_{(0,\theta)}(\max\{Y_1,Y_2\})}$

$= \dfrac{I_{(0,\theta)}(\max\{X_1,X_2\})}{I_{(0,\theta)}(\max\{Y_1,Y_2\})}$ 與 θ 無關

$\Rightarrow \max\{X_1,X_2\} = \max\{Y_1,Y_2\}$ 為必要條件.

相反地, $\max\{X_1,X_2\} = \max\{Y_1,Y_2\} \Rightarrow \dfrac{f_{X_1,X_2}(X_1,X_2)}{f_{Y_1,Y_2}(Y_1,Y_2)}$ 與 θ 無關

$\therefore \max\{X_1,X_2\}$ 為 θ 之最小充份統計量.

我們無法由 X_1+X_2 得到 $\max\{X_1,X_2\}$.

$\therefore X_1+X_2$ 並非充份統計量.

作業4

Date Page

$\boxed{10}$ $X_1, X_2, \ldots X_n \sim f(x|\theta) = \theta x^{\theta-1} I_{(0,1)}(x)$

(a) $\int_0^1 \theta x^{\theta-1} dx = 1 = \int_\infty^0 \theta (e^{-y})^{(\theta-1)} \cdot (-e^{-y}) dy$

$\left(\begin{array}{l} y = -\ln x \Leftrightarrow x = e^{-y} \\ \frac{dx}{dy} = -e^{-y} \end{array} \right.$

$= \int_0^\infty \theta e^{-\theta y + y} \cdot e^{-y} dy = \int_0^\infty \theta e^{-\theta y} dy$

由此可知, $Y \sim \exp(\theta)$ (mean: $\frac{1}{\theta}$)

(b) $V[T] \geq \frac{(\tau'(\theta))^2}{I_n(\theta)}$ ... Cramer Rao's Lower Bound

(T為$\tau(\theta)$之不偏估計量)

$I_1(\theta) = E[(\frac{\partial}{\partial\theta} \ln f(x|\theta))^2]$

$= -E[\frac{\partial^2}{\partial\theta^2} \ln f(X|\theta)] = -E[\frac{\partial^2}{\partial\theta^2}((\theta-1)\ln X + \ln\theta)]$

$= -E[-\frac{1}{\theta^2}] = \frac{1}{\theta^2}$

$\therefore I_n(\theta) = \frac{n}{\theta^2}$ $\therefore V[T] \geq \frac{\theta^2 \cdot \tau'(\theta)^2}{n}$

Cramer-Rao's Lower Bound

$$= \frac{\theta^2}{n} \cdot (\tau'(\theta))^2 \quad \tau'(\theta) = \frac{1}{\theta} \Rightarrow \frac{1}{n\theta^2}$$

(c) $Y_i = -\ln X_i \sim \exp(\theta)$ (mean $\frac{1}{\theta}$; Var $\frac{1}{\theta^2}$)

$\Rightarrow E\left[\frac{Y_1+Y_2+\cdots+Y_n}{n}\right] = \frac{1}{\theta}$ ∴ 不偏估計量

$V\left[\frac{Y_1+Y_2+\cdots+Y_n}{n}\right] = \frac{1}{n\theta^2}$ ($\because V[X_j] = \frac{1}{\theta^2}$)

↳ 達到 Cramer Rao's Lower Bound.

$\therefore \sum_{j=1}^{n} \frac{-\ln X_j}{n}$ 為 $\frac{1}{\theta}$ 之有效估計量

$\Rightarrow$ UMVUE.

[1] MLE的不變性:

(a) $L(\theta) = \prod_{i=1}^{n} f_\theta(x_i)$

$\hat{\theta}_{MLE} = \underset{\theta \in \Theta}{\arg\max}\, L(\theta)$ …… ⊛

($\theta \in \Theta$ 使得 $L(\theta)$ 最大)

$g(\theta)$ 為嚴格遞增函數。 g: 一對一函數。

$\therefore \eta = g(\theta) \Leftrightarrow \theta = g^{-1}(\eta)$

$\hat{\eta}_{MLE} = \underset{\eta}{\arg\max}\, L$

$= \underset{\eta \in g(\Theta)}{\arg\max}\, L(g^{-1}(\eta))$

註 $g(\Theta) = \{g(\theta) \mid \theta \in \Theta\}$

根據⊛, $g^{-1}(\eta) = \hat{\theta}_{MLE}$ 使得 $L(g^{-1}(\eta))$ 最大

$\therefore g^{-1}(\eta) = \hat{\theta}_{MLE} \Leftrightarrow \underbrace{g \circ g^{-1}(\eta)}_{= \eta} = g(\hat{\theta}_{MLE})$

$\therefore \hat{\eta}_{MLE} = g(\hat{\theta}_{MLE})$

(b) 接著考慮 g 不一定為一對一函數的狀況。

$$\hat{\eta}_{MLE} = \underset{\eta}{\arg\max}\, L(\theta) = \underset{\eta}{\arg\max}\, L(\beta)$$

$(\because \eta = g(\theta))$

$\begin{cases} \cdot\ \eta \in g(\Theta) \\ \cdot\ \beta \in g^{-1}(\eta) \cap \Theta = \Theta \end{cases}$

$\beta = \hat{\theta}_{MLE}$ 使得 L 最大。($\beta \in \Theta$. $\therefore$ 至少存在一個 $\hat{\theta}_{MLE}$)

假設 $\hat{\theta}_{MLE1}$ 和 $\hat{\theta}_{MLE2}$ 皆使得 L 最大。

$\{\hat{\theta}_{MLE1}, \hat{\theta}_{MLE2}\} \subseteq g^{-1}(\hat{\eta}_{MLE})$

$\hat{\theta}_{MLE1}$, $\hat{\theta}_{MLE2}$ 皆滿足 $g(\hat{\theta}_{MLE1}) = g(\hat{\theta}_{MLE2})$。

$\therefore \hat{\eta}_{MLE} = g(\hat{\theta}_{MLE1}) = g(\hat{\theta}_{MLE2})$

$\therefore$ 我們不妨將 $g(\hat{\theta}_{MLE})$ 視為 $g(\theta)$ 之 MLE.

Date Page

$\boxed{2}$ $n=5$

(a) 求最強力檢定 ⇒ 檢定力最大 ⇒ 犯下型II錯誤

之機率最低 ⇒ 要使β最小 ($\alpha=0.05$ 固定的情況下)

∴ 利用 Neyman-Pearson's lemma

$$\Lambda=\frac{f(x_1\cdots x_n\mid\lambda=1)}{f(x_1\cdots x_n\mid\lambda=2)}=\frac{x_1\exp(-x_1)\cdot x_2\exp(-x_2)\cdots x_n\exp(-x_n)}{4x_1\exp(-2x_1)\cdots 4x_n\exp(-2x_n)}$$

$$=\left(\frac{1}{4}\right)^n\exp\left((x_1+\cdots+x_n)\right)\geq k$$

充份統計量 $X_1+\cdots+X_n$ 增加 ⇒ Λ：增加

檢定函數 $\phi(X)=\begin{cases}1 & \text{if } X_1+\cdots+X_n\geq C\text{：常數}\\ 0 & \text{else}\end{cases}$

$E[\phi(X)\mid\lambda=2]=0.05$ (∵ Type I error)

$\lambda=2\Rightarrow X_j\sim\Gamma(2,\frac{1}{2})$ (mean：1)

$X_1+\cdots X_n\sim\Gamma(2n,\frac{1}{2})$

$4(X_1+\cdots+X_n)\sim\Gamma(2n,2)=\chi^2(4n)$

$Pr(X_1+X_2+\cdots+X_n\geq C\mid\lambda=2)$

$=Pr(W\geq4C)=0.05$

$$4C = \Psi_{4n}^{-1}(0.95) \quad \Psi_{4n} \cdots \chi^2_{4n} \text{ 之 cdf.}$$

$$\therefore C = \frac{1}{4}\Psi_{4n}^{-1}(0.95)$$

$$\therefore \phi(X) = \begin{cases} 1 ; & X_1 + X_2 + \cdots + X_n \geq \frac{1}{4}\Psi_{4n}^{-1}(0.95) \\ 0 ; & \text{otherwise} \end{cases}$$

(b) 檢定力 $= E[\phi(X) \mid \underbrace{\lambda = 1}_{H_1}]$

$$= \Pr(X_1 + X_2 + \cdots + X_n \geq \underbrace{\tfrac{1}{4}\Psi_{4n}^{-1}(0.95)}_{= C} \mid \lambda = 1)$$

利用中央極限定理：$\lambda = 1$ 之下，$X_1 \sim X_n \sim \Gamma(2,1)$

$$E[X_i] = 2 \quad V[X_i] = 2 < \infty$$

$$\sqrt{n}\left(\frac{\bar{X} - 2}{\sqrt{2}}\right) \xrightarrow{d} N(0,1)$$

$$\Pr(X_1 + \cdots X_n \geq C \mid \lambda = 1)$$

$$= \Pr(\bar{X} \geq \frac{C}{n} \mid \lambda = 1) = \Pr(\bar{X} - 2 \geq \frac{C}{n} - 2 \mid \lambda = 1)$$

$$= \Pr\left(\frac{\sqrt{n}}{\sqrt{2}}(\bar{X} - 2) \geq \frac{\sqrt{n}}{\sqrt{2}}\left(\frac{C}{n} - 2\right) \mid \lambda = 1\right)$$

$$= 1 - \Phi\left(\frac{\sqrt{n}}{\sqrt{2}}\left(\frac{C}{n} - 2\right)\right) \quad \text{where } C = \frac{1}{4}\Psi_{4n}^{-1}(0.95)$$

Date Page

註…(a), (b)

- $n=5$ … $C=\frac{1}{4}\psi^{-1}_{20}(0.95)=\frac{1}{4}\cdot(31.4104)=7.8526$

∴ MP-test … $\phi(X)=\begin{cases}1 \cdots & X_1+X_2+\cdots+X_n \geq 7.8526\\ 0 \cdots & \text{else}\end{cases}$

- $n=5$ … $1-\Phi\left(\frac{\sqrt{n}}{\sqrt{2}}\left(\frac{C}{n}-2\right)\right)=1-\Phi\left(\sqrt{2.5}\left(\frac{7.8526}{5}-2\right)\right)$

$=1-\Phi(-0.6791)=\Phi(0.6791)\approx\Phi(0.68)$

≈ 0.7517 … • power ≈ 0.7517

(c) YES.

考慮 $H_0: \lambda=2$ vs $H_1: \lambda=\lambda_1$

$(\lambda_1<2)$

$$\Lambda=\frac{\lambda_1^2 x_1\exp(-\lambda_1 x_1)\cdots\lambda_1^2 x_n\exp(-\lambda_1 x_n)}{4x_1\exp(-2x_1)\cdots 4x_n\exp(-2x_n)}$$

$$=\left(\frac{\lambda_1^2}{4}\right)^n\exp\left(\underbrace{(2-\lambda_1)}_{>0}(X_1+\cdots+X_n)\right)$$

∴ 充分統計量 $X_1+X_2+\cdots+X_n$: 增加 $\Rightarrow$ Λ : 增加

$\alpha=0.05$ 之下

跟(a)同理，MP 檢定 $\phi(X)=\begin{cases}1 & : X_1+X_2+\cdots X_n \geq c \\ 0 & : \text{else}\end{cases}$

$$c=\frac{1}{4}\chi^{-1}_{4n}(0.95)$$

由此可知 $\forall \lambda_1<2$、$\phi(X)$ 均為最強力檢定。

$\{\lambda_1 \mid \lambda_1<1\} \subseteq \{\lambda_1 \mid \lambda_1<2\}$

故此，對於 $H_0: \lambda=2$ vs $H_1: \lambda<1$，$\phi(X)$ 為

均勻最強力檢定。(UMP-test)

Date Page

$\boxed{3}$

(a) p：發生副作用之機率

(b) 我們欲得到的結論為 $p<0.15$

(∵我們希望新藥可以通過測試，所以想要否定 $p \geq 0.15$)

故此，$H_0: p \geq 0.15$ vs $H_1: p < 0.15$

我們先考慮對 $H_0: p=p_0$ vs $H_1: p=p_1$ $(p_0>p_1)$

之最強力檢定。(利用 Neyman-Pearson's Lemma)

$$X_i = \begin{cases} 1 & \cdots \text{第}i\text{個人發生副作用} \\ 0 & \text{〃 未發生 〃} \end{cases}$$

$(i=1 \sim n)$ $(n=400)$

$$\Lambda = \frac{\Pr(X_1=x_1, \ldots, X_n=x_n \mid p_1)}{\Pr(X_1=x_1, \ldots, X_n=x_n \mid p_0)}$$

$$= \frac{p_1^{(x_1+\cdots x_n)}(1-p_1)^{n-(x_1+\cdots x_n)}}{p_0^{(x_1+\cdots x_n)}(1-p_0)^{n-(x_1+\cdots x_n)}}$$

$$= \left(\frac{1-p_1}{1-p_0}\right)^n \cdot \left\{\frac{\left(\frac{p_1}{1-p_1}\right)}{\left(\frac{p_0}{1-p_0}\right)}\right\}^{(x_1+\cdots x_n)}$$

$$g(p) \overset{\text{def}}{=} \frac{p}{1-p} \quad (0<p<1) \cdots \text{遞增函數}$$

$$g(p_0) > g(p_1) \quad \therefore \frac{g(p_1)}{g(p_0)} < 1$$

$$\Lambda = \left(\frac{1-p_1}{1-p_0}\right)^n \underbrace{\left(\frac{g(p_1)}{g(p_0)}\right)}_{<1}^{(X_1+X_2+\cdots+X_n)} > k \Rightarrow \text{棄却}$$

$\therefore$ 充份統計量 $X_1+X_2+\cdots+X_n$：增 $\Rightarrow$ Λ：減少

$\therefore$ 最強力檢定 $\varphi(X) = \begin{cases} 1 : & X_1+\cdots+X_n < C \\ r : & X_1+\cdots+X_n = C \quad (\text{隨機化}) \\ 0 : & X_1+\cdots+X_n > C \end{cases}$

$(E[\varphi(X) \mid p_0] = \alpha)$

接著考慮隨機化檢定 $\varphi^* = \begin{cases} 1 & (\text{with probability } \alpha) \\ 0 & (1-\alpha) \end{cases}$

由於 $\varphi(X)$ 為最強力檢定，故滿足：

$$E[\varphi(X) \mid p_1] \geqq E[\varphi^*(X) \mid p_1] = \alpha = E[\varphi(X) \mid p_0]$$

由此可知，$\forall p_0, p_1 \ (p_0<p_1)$ $\beta_\varphi(p_0) \leqq \beta_\varphi(p_1) \cdots \circledast$

Date Page

根據先前的討論，對於 $H_0: p=p_0=0.15$

vs $H_1: p=p_1\ (<0.15)$ 顯著水準 α 下的

最強力檢定亦為 對於 $H_0: p\geq 0.15$

vs $H_1: p<0.15$ 顯著水準 α 下的 均勻最強力

檢定.

- $\beta_\varphi(p)$ 為遞減函數，故 $\sup_{p\geq 0.15}\beta_\varphi(p)=\beta_\varphi(0.15)$

$\Rightarrow$ 顯著水準不變.

- $\forall p_1 > p_0 = 0.15$ 滿足 $\beta_\varphi(p_1)\geq\beta_{\varphi^*}(p_1)$

(β_{φ^*} 為 任意顯著水準為 α 的檢定之檢定力)

($\because$ 對立假設變成 $H_0: p<0.15$，φ 依然為

最強力 (for all $p_1<0.15$)

接著求 C 使得 $E[\varphi(X)\mid p=0.15]=\alpha$

$=Pr(X_1+X_2+\cdots X_n\leq C\mid p=0.15)=\alpha$

利用中央極限定理. ($n=400$, $p=0.15$)

$$\frac{X_1+X_2+\cdots X_n}{n}\xrightarrow{d}N\left(p,\frac{p(1-p)}{n}\right)$$

$$Z \overset{def}{=} \frac{\sqrt{400}(\bar{X}-0.15)}{\sqrt{0.85}\sqrt{0.15}} \qquad Pr\left(Z \leq \frac{\sqrt{400}\left(\frac{c}{n}-0.15\right)}{\sqrt{0.85}\sqrt{0.15}}\right) = \alpha$$

$(n=400)$

$$\therefore C = 60 + 20\sqrt{0.85}\cdot\sqrt{0.15}\cdot\Phi^{-1}(\alpha)$$

$$\therefore \varphi(X) = \begin{cases} 1 \cdots & \text{if } X_1+X_2+\cdots X_{400} \leq 60+20\sqrt{0.85}\sqrt{0.15}\,\Phi^{-1}(\alpha) \\ 0 \cdots & \text{otherwise} \end{cases}$$

為對 $H_0: p \geq 0.15$ vs $H_1: p < 0.15$ 之 UMP 檢定.

(c) $P(X_1+X_2+\cdots X_{400} \leq 80 \mid p=0.15) = P\left(Z \leq \frac{1}{\sqrt{0.85}\cdot\sqrt{0.15}}\right)$

$\therefore$ p 值 $= \Phi\left(\frac{1}{\sqrt{0.85}\sqrt{0.15}}\right)$

(d) $\alpha = 0.05$ 時...

$$\varphi(X) = \begin{cases} 1 & \text{if } X_1+X_2+\cdots X_{400} \leq 71.748 \\ 0 & \text{else} \end{cases}$$

$X_1+X_2+\cdots+X_{400} = 80 \Rightarrow \varphi(X) = 0$

$\therefore$ 不棄卻 H_0.

Date Page

4 利用 χ^2-test.

$\begin{cases} H_0: \text{資料服從 Poisson 分佈} \\ H_1: \text{資料不服從 Poisson 分佈} \end{cases}$

檢定統計量 $= \sum_{i=0}^{k} \frac{(E_i - O_i)^2}{E_i}$

H_0 之下的分佈 ∴ χ^2_{k-1}

(∵ 自由度 $= (k-0+1)-1-\underset{\sim}{1}$ 估計的參數的個數 (λ))

① 求 λ 的 MLE

$$\hat{\lambda}_{MLE} = \bar{X} = \frac{0\cdot229+1\cdot211+2\cdot93+3\cdot35+4\cdot7+5\cdot1}{576}$$

$$= \frac{211+186+105+28+5}{576}$$

$$= \frac{535}{576} = 0.9288$$

② $E_i = 576 \cdot P_r(X_1 = i) \quad (i = 0 \sim 4)$

$E_i = 576 \cdot P_r(X_1 \geq i) \quad (i = 5)$

(∵ $k = 5$)

$E_0 = 227.5314$

$E_1 = 211.3356$

$E_2 = 98.1463$

$E_3 = 30.3867$

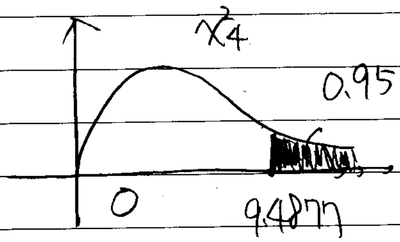
$E_4 = 7.0559$

$E_5 = 1.5441$

③ 檢定統計量 $T = \sum_{i=0}^{k} \frac{(E_i - O_i)^2}{E_i}$

$= 0.9807.$

$T \sim \chi^2_4$



$\overset{T}{0.9807} \ll 9.4877$

∴ 不棄卻 H_0

（無法否定資料服從 Poisson 分佈）

Date Page

5 假設資料服從常態分佈且舊的壽命

與新的壽命之變異數是共同的.

$\therefore X_1 \sim X_n \sim N(\mu, 1.1^2)$（新的壽命之資料）

我們欲檢定 $H_0: \mu = 4.2$ vs $H_1: \mu > 4.2$

我們先利用 Neyman-Pearson's lemma 來求

對於 $H_0: \mu = \mu_0 = 4.2$ $H_1: \mu = \mu_1$ $(\mu_1 > \mu_0)$

之最強力檢定.

$$\Lambda(X) \overset{def}{=} \frac{f(x|\mu_1)}{f(x|\mu_0)} = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_1)^2\right)}{\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right)}$$

（注：$\sigma^2 = 1.1^2$；$n = 80$）

$$= \exp\left(\frac{1}{\sigma^2}\underbrace{(\mu_1-\mu_0)}_{>0}(X_1+X_2+\cdots X_n) - \frac{n\mu_1^2}{2\sigma^2} + \frac{n\mu_0^2}{2\sigma^2}\right)$$

$X_1+X_2+\cdots X_n \cdots \mu$ 之充份統計量

$X_1+X_2+\cdots X_n$：增加 $\Leftrightarrow$ Λ：增加

$$\therefore \varphi(X) = \begin{cases} 1 & \text{if } X_1+X_2+\cdots X_n \geq C \\ 0 & \text{else} \end{cases}$$

設顯著水準為α， $E[\varphi(X) \mid \mu=\mu_0]=\alpha$

$\mu=\mu_0 \Rightarrow \sqrt{n}\left(\frac{\bar{X}-\mu_0}{\sigma}\right) \sim N(0,1)$

$Pr\left(\frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma} \geq \frac{\sqrt{n}(\frac{c}{n}-\mu_0)}{\sigma}\right)=1-\Phi\left(\frac{\sqrt{n}(\frac{c}{n}-\mu_0)}{\sigma}\right)=\alpha$

$\therefore \frac{\sqrt{n}}{\sigma}\left(\frac{c}{n}-\mu_0\right)=\Phi^{-1}(1-\alpha)$

$\therefore \frac{c}{n}-\mu_0=\frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\alpha) \quad \therefore c=n\mu_0+\sigma\sqrt{n}\,\Phi^{-1}(1-\alpha)$

$\therefore \varphi(X)=\begin{cases} 1 & : \text{if } \bar{X} \geq \mu_0+\frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\alpha) \\ 0 & : \text{else} \end{cases}$

$\forall \mu_1>\mu_0$， $\varphi(X)$均為顯著水準α下的最強力檢定

對於 $H_0: \mu=\mu_0=4.2$ vs $H_1: \mu>\mu_0$，

$\varphi(X)$為顯著水準α下的均勻最強力檢定

$C_\alpha=\left\{(X_1,\dots,X_n) \mid \bar{X} \geq \mu_0+\frac{\sigma}{\sqrt{n}}\Phi^{-1}(1-\alpha)\right\}$

$\text{p-value}=\inf_{\alpha} (X_1,\dots,X_n)\mid_{\bar{X}=4.5} \in C_\alpha$

$4.5=4.2+\frac{1.1}{\sqrt{80}}\Phi^{-1}(1-\alpha) \Rightarrow \Phi^{-1}(1-\alpha)=2.44$

$\Rightarrow 1-\alpha=0.9927 \Rightarrow \alpha=0.0073 \quad \therefore \text{p-value}=0.0073$

由於p-value偏小，故棄卻H_0。

Date Page

6

- 虛無假設…試著推翻的假設
- 對立假設…試著證明的假設
  (具意外性的假設)

在信任員工的情況下，「某某侵占公款」被視為比「某某並未侵占公款」(=無辜)更具意外性的消息。

故此，

- H_0：〇〇無辜的
- H_1：〇〇是兇手.

(a) 型Ⅰ錯誤…
雖然 H_0 為真，卻棄卻 H_0 的錯誤
⇔ 冤枉無辜店員的錯誤

(b) 型II錯誤…

雖然 H_1 為真，卻未拒卻 H_0 之錯誤

(二) 放過兇手的錯誤。

* 犯下型I錯誤通常比較嚴重。

· 照這個案例來說，在信任自己員工的情況下，你會覺得冤枉無辜的員工是一個很嚴重的錯誤。(比放過兇手更嚴重)

· 相反地，如果你不信任自己員工的話，H_0 跟 H_1 會顛倒。(H_0：OO是兇手；H_1：OO無辜)

在這種情況下，保護公司資產比員工名譽更重要，所以你會認為放過兇手(=型I錯誤)比較嚴重。

⇒ 無論 H_0，H_1 為何，犯下型I錯誤比較嚴重。

$\boxed{7}$

(a) $\Phi_n(\theta) = \frac{1}{\theta} - \overline{X} = 0 \Leftrightarrow \theta = \frac{1}{\overline{X}}$ $(x_j > 0)$

(b) $X_j \sim \exp(\theta)$ (mean $\frac{1}{\theta}$)

$Y_j \overset{\text{def}}{=} \frac{1}{\theta} - X_j$ … (mean 0)

根據弱大數法則 $\frac{Y_1 + Y_2 + \cdots + Y_n}{n} \xrightarrow{P} 0$

$\therefore \frac{1}{\theta} - \overline{X} \xrightarrow{P} 0$

$\therefore \Phi_n(\theta) \xrightarrow{P} 0$

(c) $\hat{\theta}_n = \frac{1}{\overline{X}}$

$g(t) = \frac{1}{t}$ $(t > 0)$ 為連續函數.

根據弱大數法則, $\overline{X} \xrightarrow{P} \theta$.

$g(\overline{X}) \xrightarrow{P} g(\theta)$ $\therefore \frac{1}{\overline{X}} \xrightarrow{P} \frac{1}{\theta}$

(d) 根據中央極限定理,

$\sqrt{n}(\overline{X} - \frac{1}{\theta}) \xrightarrow{d} N(0, \frac{1}{\theta^2})$

利用 δ-method …

$\sqrt{n}(g(\overline{X}) - g(\frac{1}{\theta})) \xrightarrow{d} N(0, \frac{1}{\theta^2}\{g'(\frac{1}{\theta})\}^2)$

$\therefore \sqrt{n}\left(\frac{1}{\bar{X}}-\theta\right) \xrightarrow{d} N(0,\theta^2) \quad (\theta>0)$

(e) $L(\theta|x) = \prod_{j=1}^{n} f(x_j|\theta) = \theta^n \exp\left(-\theta(x_1+x_2+\cdots+x_n)\right)$

$\log L(\theta|x) = n\log\theta - \theta(x_1+x_2+\cdots+x_n) = 0$

$\frac{\partial}{\partial\theta}\log L(\theta|x) = \frac{n}{\theta} - (x_1+\cdots x_n) = 0 \quad \therefore \hat{\theta}_{MLE} = \frac{1}{\bar{x}}$

($\hat{\theta}_{MLE}$ 使得 L 最大)

$\hat{\theta}_{MLE} = \hat{\theta}_n$

根據 (d) $\sqrt{n}(\hat{\theta}_{MLE}-\theta) \xrightarrow{d} N(0,\theta^2) = N(0, I(\theta)^{-1})$

$\therefore I(\theta) = \frac{1}{\theta^2}$ (觀測到一個樣本時得到的 Fisher 資訊量)

Date Page

$\boxed{8}$ $X_1, X_2, \dots X_n \sim Po(\theta)$ $(\theta > 0)$

$n\bar{X} = X_1 + X_2 + \dots X_n \sim Po(n\theta)$

$Y \overset{def}{=} X_1 + X_2 + \dots + X_n$

$E\left[\left(1-\frac{a}{n}\right)^Y\right] = \sum_{y=0}^{\infty} \left(1-\frac{a}{n}\right)^y \cdot e^{-n\theta} \cdot \frac{(n\theta)^y}{y!}$

$= e^{-n\theta} \sum_{y=0}^{\infty} \frac{(n\theta - a\theta)^y}{y!} = e^{-n\theta} \cdot \sum_{y=0}^{\infty} \frac{\{\theta(n-a)\}^y}{y!}$

$= e^{-n\theta} \exp(n\theta - \theta a) = \exp(-\theta a)$

由此可知, $\left(1-\frac{a}{n}\right)^Y$ 為 $\exp(-a\theta)$ 之不偏估計量

$\therefore MSE\left(\left(1-\frac{a}{n}\right)^Y ; \exp(-a\theta)\right) = V\left[\left(1-\frac{a}{n}\right)^Y\right]$

$= E\left[\left(1-\frac{a}{n}\right)^{2Y}\right] - \underbrace{E\left[\left(1-\frac{a}{n}\right)^Y\right]^2}_{\exp(-2a\theta)}$

$\sum_{y=0}^{\infty} \left(1-\frac{a}{n}\right)^{2y} e^{-n\theta} \cdot \frac{(n\theta)^y}{y!}$

$= e^{-n\theta} \sum_{y=0}^{\infty} \left(1 - \frac{2a}{n} + \frac{a^2}{n^2}\right)^y \cdot (n\theta)^y \cdot \frac{1}{y!}$

$= e^{-n\theta} \sum_{y=0}^{\infty} \left\{\left(n - 2a + \frac{a^2}{n}\right)\theta\right\}^y \cdot \frac{1}{y!}$

$$= e^{-n\theta} \cdot \exp\left(n\theta - 2a\theta + \frac{a^2\theta}{n}\right) = \exp\left(-2a\theta + \frac{a^2\theta}{n}\right)$$

$$\therefore V\left[\left(1-\frac{a}{n}\right)^{Y}\right] = \exp\left(-2a\theta + \frac{a^2\theta}{n}\right) - \exp(-2a\theta)$$

$\parallel$

(MSE)

bias $= 0$

Date Page

$\boxed{9}$ $X_1, X_2 \cdots X_n \sim N(\theta, \theta)$

(1) $X_1+X_2+\cdots+X_n \sim N(n\theta, n\theta)$

$\bar{X} \sim N(\theta, \frac{\theta}{n}) \Rightarrow \frac{\sqrt{n}}{\sqrt{\theta}}(\bar{X}-\theta) \sim N(0,1)$

(2) 利用 δ-method:

$g(x) \overset{def}{=} \cos x$

$\frac{\sqrt{n}}{\sqrt{\theta}}(g(\bar{X})-g(\theta)) \xrightarrow{d} N(0, \{g'(\theta)\}^2)$

$g'(x) = -\sin x \quad \{g'(x)\}^2 = \sin^2 x$

$\therefore \frac{\sqrt{n}}{\sqrt{\theta}}(\cos\bar{X}-\cos\theta) \xrightarrow{d} N(0, \sin^2\theta)$

$\therefore \sqrt{n}(\cos\bar{X}-\cos\theta) \xrightarrow{d} N(0, \theta\sin^2\theta)$

若 $\theta = k\pi$ $(k=1,2,3\cdots)$ 時, $\sqrt{n}(\cos\bar{X}-\cos\theta)$

之變異數為 0. $\Rightarrow$ 退化.

Date Page

$\boxed{10}$ $X \sim N(\theta, 1)$

$H_0: \theta \in \Theta_0 = \{0\}$ vs $H_1: \theta \in \Theta_1 = \{1, -1\}$

證明不存在UMP檢定

先考慮 ① $H_0: \theta = 0$ vs $H_1: \theta = 1$

② $H_0: \theta = 0$ vs $H_1: \theta = -1$

① 利用 Neyman-Pearson's lemma:

$$\Lambda(X) = \frac{\prod_{j=1}^{n} f(x_j \mid \theta = 1)}{\prod_{j=1}^{n} f(x_j \mid \theta = 0)} = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{j=1}^{n}(x_j - 1)^2\right)}{\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2}\sum_{j=1}^{n}(x_j)^2\right)}$$

$$= \exp\left(\sum_{j=1}^{n} x_j - \frac{n}{2}\right) \geq k_1 \Leftrightarrow \sum_{j=1}^{n} x_j \geq c_1$$

($\sum_{j=1}^{n} x_j$ 為 θ之充份統計量)

$$\varphi_1(X) = \begin{cases} 1 \cdots & \sum_{j=1}^{n} x_j \geq c_1 \\ 0 \cdots & \text{elsewhere} \end{cases}$$

顯著水準: $E[\varphi_1(X) \mid \theta = 0] = \alpha$

$= 1 - \Phi\left(\frac{c_1}{\sqrt{n}}\right)$ $\therefore$ $c_1 = \sqrt{n}\,\Phi^{-1}(1-\alpha)$

$$\therefore \varphi_1(X) = \begin{cases} 1 \cdots & \sum_{j=1}^{n} x_j \geq \sqrt{n}\,\Phi^{-1}(1-\alpha) \\ 0 & \text{otherwise} \end{cases}$$

為對 $H_0: \theta = 0$ vs $H_1: \theta = 1$ 之MP檢定.

② $H_0: \theta=0$ vs $H_1: \theta=-1$

同樣道理、$\varphi_2(X)=\begin{cases} 1 \cdots & \sum_{i=1}^{n} X_i \leq \sqrt{n}\,\Phi^{-1}(\alpha) \\ 0 \cdots & \text{otherwise} \end{cases}$

為 MP 檢定

③ 接著考慮 對於「$H_0: \theta=0$ vs $H_1: \theta\in\{1,-1\}$」之

UMP 檢定 φ. $\left(\begin{array}{l} \beta_1(\theta) \overset{def}{=} E[\varphi_1|\theta],\ \beta_2(\theta) \overset{def}{=} E[\varphi_2|\theta] \\ \beta(\theta) = E[\varphi|\theta] \end{array}\right)$

$\beta(\theta) = \max\{\beta_1(\theta), \beta_2(\theta)\}$ $(\forall\theta\in\{1,-1\})$

$\theta=1$ 時 $\beta(1)=\beta_1(1)$ $\therefore \varphi=\varphi_1$

$\theta=-1$ 時 $\beta(-1)=\beta_2(-1)$ $\therefore \varphi=\varphi_2$

$\therefore$ 若存在 UMP 檢定，則需滿足 $\varphi_1=\varphi_2$ (a.s).

$\therefore$ 但 $\varphi_1\neq\varphi_2$ $\therefore$ 不存在 UMP 檢定

Date Page

$\boxed{III}$

① 先考慮對 $H_0: \theta=0$ vs $H_1: \theta=1$ 之最強力檢定
(顯著水準 α)

利用 Neyman Pearson's lemma：

$$\Lambda(X) = \frac{\prod_{j=1}^{n} f(x_j \mid \theta=1)}{\prod_{j=1}^{n} f(x_j \mid \theta=0)} = \exp\left((X_1+X_2+\cdots+X_n) - \frac{n}{2}\right) \geq k_1$$

$$\varphi_1(X) = \begin{cases} 1 & \cdots \text{ if } X_1+X_2+\cdots X_n \geq C_1 \\ 0 & \cdots \text{ else} \end{cases}$$

$$E[\varphi_1(X) \mid \theta=0] = \alpha$$

$$\parallel$$

$$1-\Phi\left(\frac{C_1}{\sqrt{n}}\right) = \alpha \quad \therefore \quad C_1 = \sqrt{n}\,\Phi^{-1}(1-\alpha)$$

$$\therefore \varphi_1(X) = \begin{cases} 1 & X_1+X_2+\cdots X_n \geq \sqrt{n}\,\Phi^{-1}(1-\alpha) \\ 0 & \text{elsewhere} \end{cases}$$

② 接著考慮 $H_0: \theta=0$ vs $H_1: \theta=2$ 之情形.

$$\Lambda(X) = \frac{\prod_{j=1}^{n} f(x_j \mid \theta=2)}{\prod_{j=1}^{n} f(x_j \mid \theta=0)} \geq k_2$$

$$\parallel$$

$$\exp\left(2(X_1+X_2+\cdots+X_n) - 2n\right)$$

$$\therefore \varphi_2(X) = \begin{cases} 1 & \text{if } X_1+X_2+\cdots X_n \geq C_2 \\ 0 & \text{else} \end{cases}$$

$E[\varphi_2(X) \mid \theta=0] = \alpha$

$\therefore C_2 = \sqrt{n}\,\Phi^{-1}(1-\alpha)$

$$\varphi_2(X) = \begin{cases} 1 & \text{if } X_1+X_2+\cdots+X_n \geq \sqrt{n}\,\Phi^{-1}(1-\alpha) \\ 0 & \text{elsewhere} \end{cases}$$

由此可知 $\varphi_1(X)=\varphi_2(X)$.

故此無論 $\theta=1$ 還是 $\theta=2$, $\varphi_1(X)$ $(\varphi_2(X))$ 均擁有最大的檢定力.

$\therefore$ $\varphi_1(X)$ $(=\varphi_2(X))$ 為對於 $H_0: \theta=0$ vs $H_1: \theta\in\{1,2\}$ 顯著水準 α 下的均勻最強力檢定.

Date Page (a) 題目應為 $\ell(\hat{\theta})-\ell(\theta)$ is increasing or decreasing

$(\ell = \ln f)$

$\boxed{12}$ $L \overset{def}{=} f(x|\theta)$ $\ell \overset{def}{=} \log L$

(a) $\log f(x|\theta) = \eta(\theta)\,Y(x) - \xi(\theta) + \log h(x) = \ell(\theta)$

$\frac{\partial}{\partial\theta}\log f(x|\theta) = \frac{\partial \ell}{\partial\theta} = \eta'(\theta)\,Y(x) - \xi'(\theta) = 0$

$\hat{\theta}_{MLE}\ (= \hat{\theta}(X))$ 滿足 $\eta'(\hat{\theta})\,Y(x) - \xi'(\hat{\theta}) = 0$.

$$\frac{d}{dY}\left(\ell(\hat{\theta}) - \ell(\theta_0)\right) = \frac{dx}{dY}\cdot\frac{d}{dx}\underbrace{\left(\ell(\hat{\theta}) - \ell(\theta_0)\right)}_{\left(\eta(\hat{\theta})Y(x) - \xi(\hat{\theta}) - \eta(\theta_0)Y(x) + \xi(\theta_0)\right)}$$

$$= \left(\frac{dY}{dx}\right)^{-1}\cdot\left\{\eta'(\hat{\theta})\cdot\hat{\theta}'(x)\,Y(x) + \eta(\hat{\theta})\,Y'(x) - \xi'(\hat{\theta})\cdot\hat{\theta}'(x) - \eta(\theta_0)\,Y'(x)\right\}$$

$$= \left(\frac{1}{Y'(x)}\right)\cdot\left\{\hat{\theta}'(x)\underbrace{\left(\eta'(\hat{\theta})\,Y(x) - \xi'(\hat{\theta})\right)}_{=0} + Y'(x)\left(\eta(\hat{\theta}) - \eta(\theta_0)\right)\right\}$$

$= \eta(\hat{\theta}) - \eta(\theta_0)$

由於 η 為遞增函數，故 $\eta(\hat{\theta}) - \eta(\theta_0) \geq 0$ $(\hat{\theta} \geq \theta_0)$

$\therefore \frac{d}{dY}\left(\ell(\hat{\theta}) - \ell(\theta_0)\right) \geq 0$

由此可知 $\ell(\hat{\theta}) - \ell(\theta)$ is increasing in Y $(\hat{\theta} \geq \theta_0)$

$\therefore$ 證明完成

(b) 在討論概似比檢定前，我們有必要

證明 $\frac{d\theta}{dY} \geq 0$。（假設 $\frac{d}{d\theta}$ 與 $\int$ 可替換）

• 全機率 $=1$ ∴ $\int_R f(x|\theta)\,dx = 1$

$\Rightarrow \int_R \frac{d}{d\theta} f(x|\theta)\,dx = 0$

$\Rightarrow \int_R \left(\eta'(\theta) Y(x) - \xi'(\theta)\right) f(x|\theta)\,dx = 0$

由此可知 $E[Y(x)] = \frac{\xi'(\theta)}{\eta'(\theta)}$

• 同樣道理，$\int_R \{\eta''(\theta) Y(x) - \xi''(\theta)\} f(x|\theta)\,dx$

$+ \int_R \left(\eta'(\theta) Y(x) - \xi'(\theta)\right)^2 f(x|\theta)\,dx = 0$

$\Rightarrow \eta''(\theta) \cdot \frac{\xi'(\theta)}{\eta'(\theta)} - \xi''(\theta) + \eta'(\theta)^2 E[Y(x)^2] - 2\eta'(\theta)\xi'(\theta)\frac{\xi'(\theta)}{\eta'(\theta)} + (\xi'(\theta))^2 = 0$

$\Leftrightarrow \eta''(\theta)\frac{\xi'(\theta)}{\eta'(\theta)} - \xi''(\theta) + \eta'(\theta)^2 E[Y^2] - (\xi'(\theta))^2 = 0$

$\Leftrightarrow \eta'(\theta)^2 E[Y^2] = \xi''(\theta) + \xi'(\theta)^2 - \eta''(\theta) \cdot \frac{\xi'(\theta)}{\eta'(\theta)}$

$\therefore E[Y^2] = \frac{\xi''(\theta)}{\eta'(\theta)^2} + \underbrace{\left(\frac{\xi'(\theta)}{\eta'(\theta)}\right)^2}_{= E[Y]^2} - \eta''(\theta) \cdot \frac{\xi'(\theta)}{\eta'(\theta)^3}$

Date Page

$$\therefore V[Y] = \frac{\zeta''(\theta)}{(\eta'(\theta))^2} - \eta''(\theta)\cdot\frac{\zeta'(\theta)}{\eta'(\theta)^3}$$

$$= \frac{\zeta''(\theta)\eta'(\theta) - \eta''(\theta)\zeta'(\theta)}{(\eta'(\theta))^3} > 0$$

$\eta'(\theta) \geq 0$ （$\because \eta(\theta)$ 為廣增函數

$\therefore \zeta''(\theta)\eta'(\theta) - \eta''(\theta)\zeta'(\theta) > 0$ …… ⊛

・$\hat{\theta}_{MLE}$ 之定理…… $\eta'(\hat{\theta})T(X) - \zeta'(\hat{\theta}) = 0$

$$\therefore T(X) = \frac{\zeta'(\hat{\theta})}{\eta'(\hat{\theta})}$$

$$\therefore \frac{dT}{d\hat{\theta}} = \frac{\zeta''(\hat{\theta})\eta'(\hat{\theta}) - \zeta'(\hat{\theta})\eta''(\hat{\theta})}{(\eta'(\hat{\theta}))^2} > 0 \quad (\because ⊛)$$

$\left(\frac{d\hat{\theta}}{dT} > 0\right)$ $\therefore$ T: 增加 $\Leftrightarrow$ $\hat{\theta}$: 增加

接著考慮概似比検定（$H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$）

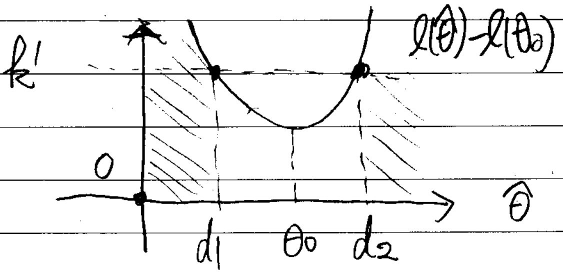
$$\Lambda(X) = \frac{\sup_{\theta=\theta_0} f(X|\theta)}{\sup_{\theta\in\Theta} f(X|\theta)} = \frac{f(X|\theta_0)}{\sup_{\theta\in\Theta} f(X|\theta)} < k \Rightarrow \text{棄卻}$$

$$-\ln\Lambda(X) = \ell(\hat{\theta}) - \ell(\theta_0) > k' \Rightarrow \text{棄卻}$$

（註 $\ell = \ln f$

$\because \hat{\theta} > \theta_0$ 時… $\hat{\theta}$：增加：$\ell(\hat{\theta}) - \ell(\theta_0)$：增加

$\hat{\theta} < \theta_0$ 時… $\hat{\theta}$：增加：$\ell(\hat{\theta}) - \ell(\theta_0)$：減少



$\therefore$ 棄卻域：$\{\hat{\theta} \mid \hat{\theta} < d_1$ or $\hat{\theta} > d_2\}$

$\Leftrightarrow \{Y \mid Y < C_1$ or $Y > C_2\}$

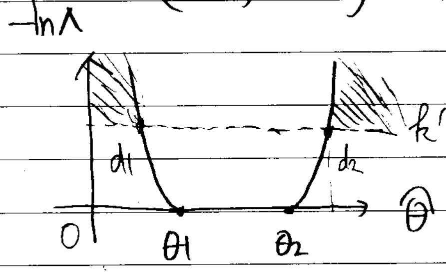
$(\because \frac{dY}{d\hat{\theta}} > 0)$

$\therefore$ 證明完成. ∎

(三) $H_0: \theta \in [\theta_1, \theta_2]$ vs $H_1: \theta \notin [\theta_1, \theta_2]$ 的 case.

$$\sup_{\theta \in [\theta_1, \theta_2]} \ln f(X|\theta) = \begin{cases} \ell(\theta_1) & (\hat{\theta} < \theta_1) \\ \ell(\hat{\theta}) & (\theta_1 \leq \hat{\theta} \leq \theta_2) \\ \ell(\theta_2) & (\hat{\theta} > \theta_2) \end{cases}$$

$$\therefore -\ln \Lambda = \begin{cases} \ell(\hat{\theta}) - \ell(\theta_1) & (\hat{\theta} < \theta_1) & \cdots \text{(減少函數)} \\ 0 & (\theta_1 \leq \hat{\theta} \leq \theta_2) & \cdots (0) \\ \ell(\hat{\theta}) - \ell(\theta_2) & (\hat{\theta} > \theta_2) & \text{(增加函數)} \end{cases}$$



$\therefore \hat{\theta} < d_1$ or $\hat{\theta} > d_2$

$\Leftrightarrow Y < C_1$ or $Y > C_2$

$\therefore$ 證明完成. ∎

$\boxed{13}$ 假設 F, G 為可微函數.

(1) 令 $h(x) = \frac{d}{dx}\left(\theta F(x) + (1-\theta) G(x)\right) = \theta f(x) + (1-\theta) g(x)$

先考慮對 $H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$ $(\theta_0 < \theta_1)$

之最強力檢定. (令其顯著水準 $= \alpha$)

$$\Lambda(x) = \frac{h(x|\theta_1)}{h(x|\theta_0)} = \frac{\theta_1 f(x) + (1-\theta_1) g(x)}{\theta_0 f(x) + (1-\theta_0) g(x)} \geq k$$

$$\Leftrightarrow \frac{\theta_1 \cdot \frac{f(x)}{g(x)} + (1-\theta_1)}{\theta_0 \frac{f(x)}{g(x)} + (1-\theta_0)} \geq k$$

$$u(x) \overset{\text{def}}{=} \frac{f(x)}{g(x)}$$

$$\Lambda(x) = \frac{\theta_1 u(x) + (1-\theta_1)}{\theta_0 u(x) + (1-\theta_0)} = \frac{\theta_1}{\theta_0} + \frac{\left(1 - \frac{\theta_1}{\theta_0}\right)}{\theta_0 u(x) + (1-\theta_0)}$$

$$= \frac{\theta_1}{\theta_0} - \frac{\left(\frac{\theta_1}{\theta_0} - 1\right)}{\theta_0 u(x) + (1-\theta_0)} \geq k$$

$$\Leftrightarrow \left(\frac{\theta_1}{\theta_0} - k\right) \geq \frac{\left(\frac{\theta_1}{\theta_0} - 1\right)}{\theta_0 u(x) + (1-\theta_0)} \qquad \left(\frac{\theta_1}{\theta_0} - 1 > 0\right)$$

u 增加 $\Rightarrow \dfrac{\frac{\theta_1}{\theta_0} - 1}{\theta_0 u(x) + (1-\theta_0)}$ … 減少

$$\varphi(X) \overset{def}{=} \begin{cases} 1 \cdots & X \in C \\ 0 \cdots & X \notin C \end{cases} \qquad C \overset{def}{=} \{X \mid U(X) \geq C_\alpha\}$$

$$E[\varphi(X) \mid \theta_0] = \alpha$$

$\varphi(X)$為對 $H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$，顯著水準α下的最強力檢定，

我們另外考慮 $\varphi^*(X) = \begin{cases} 1 & (\text{with probability } \alpha) \\ 0 & (\quad " \quad 1-\alpha) \end{cases}$

（隨機化檢定）

$\beta(\theta) \overset{def}{=} E[\varphi(X) \mid \theta]$，$\beta^*(\theta) \overset{def}{=} E[\varphi^*(X) \mid \theta] = \alpha \Rightarrow$

$\beta(\theta_1) \geq \beta^*(\theta_1) = \alpha = \beta(\theta_0)$ $\quad \therefore \forall \theta_1, \theta_0 \quad \beta(\theta_1) \geq \beta(\theta_0)$

由此可知，$\beta(\theta)$為遞增函數。 $\therefore \sup\limits_{\theta \leq \theta_0} \beta(\theta) = \alpha$

另外，$\forall \theta > \theta_0$，$\varphi(X)$均擁有最強力。

$\therefore$ 對於 $H_0: \theta \leq \theta_0$ vs $H_1: \theta > \theta_0$，

$\varphi(X)$為顯著水準α下的均勻最強力檢定。

（= Karlin Rubin 定理）

(2) 令$\tau(X)$為顯著水準α下對 $H_0: \theta \in \Theta_0$

vs $H_1: \theta \in \Theta_1$ 之UMP檢定.

$$\begin{cases} \Theta_0 = [0, \theta_1] \cup [\theta_2, 1] \\ \Theta_1 = (\theta_1, \theta_2) \end{cases}$$

$$\beta_\tau(\theta) = E[\varphi(X) \mid \theta]$$

① τ為非隨機化檢定之 case

$$\tau(X) = \begin{cases} 1 & X \in C_\tau \\ 0 & X \notin C_\tau \end{cases}$$

$$\beta_\tau(\theta) = \int_{C_\tau} (\theta f(x) + (1-\theta) g(x))\, dx$$

$$= \int_{C_\tau} g(x)\, dx + \theta \int_{C_\tau} (f-g)\, dx$$

由於 τ為UMP檢定，故應滿足：

$$\sup_{\theta \in \Theta_0} \beta(\theta) = \alpha \leq \beta(\theta) \quad (\theta \in \Theta_1)$$

↳ 由此可知 θ之係數必定為 0.

$$\begin{cases} \therefore \int_{C_\tau} (f-g)\, dx = 0 \\ \Rightarrow \int_{C_\tau} g(x)\, dx = \alpha \end{cases}$$

$\therefore \beta(\theta)=\alpha$ (for all θ)

由此可知，UMP檢定的檢定力，無論θ均為相同。

但是無法保證隨時都存在C_τ使得

$\int_{C_\tau} g(x)\,dx = \int_{C_\tau} f(x)\,dx = \alpha$.

但τ之檢定力與顯著水準α下的隨機化檢定相同。$\therefore$我們直接考慮隨機化檢定即可。($\Rightarrow$ ②)

② τ為隨機化檢定的case

$$\tau(x) = \begin{cases} 1 & (\text{with probability } \alpha) \\ 0 & (\quad " \quad 1-\alpha) \end{cases} = T(x)$$

根據①的討論，$\tau(x)$ $(=T(x))$為顯著水準α下的UMP檢定。

Date Page

(3) 對於 $H_0: \theta \le \theta_1$ or $\theta \ge \theta_2$ vs $H_1: \theta_1 < \theta < \theta_2$

之概似比檢定。

$$\Lambda(x) = \frac{\sup\limits_{\theta \le \theta_1 \text{ or } \theta \ge \theta_2} f(x|\theta)}{\sup\limits_{\theta \in [0,1]} f(x|\theta)} < \lambda$$

$$\frac{\partial f(x|\theta)}{\partial \theta} = f(x) - g(x) \quad (\text{與 } \theta \text{ 無關})$$

① $f(x) - g(x) \ge 0$ 的 case:

$L(\theta|x) = f(x|\theta)$ 為 θ 之 遞增函數,

∴ $\theta = 1$ 使得 $L(\theta|x) = f(x|\theta)$ 最大.

② $f(x) - g(x) < 0$ 的 case

$L(\theta|x) = f(x|\theta)$ 為 θ 之 遞減函數.

$\theta = 0$ 使得 $L(\theta|x) = f(x|\theta)$ 最大.

無論 ① 或 ②, $\sup\limits_{\substack{\theta \le \theta_1 \text{ or} \\ \theta \ge \theta_2}} f(x|\theta) \Big/ \sup\limits_{\theta \in [0,1]} f(x|\theta) = 1$

$\therefore \Lambda(X) = 1 \quad (\text{for all } X)$

$\therefore$ 利用随機化檢定，

$$\varphi_\Lambda(X) = \begin{cases} 1 & (\text{with probability } \alpha) \\ 0 & (\quad " \quad 1-\alpha) \end{cases}$$

Date Page

$\boxed{14}$ (1) $\theta_1 > \theta_0$

我們希望證明 $\dfrac{f(x|\theta_1)}{f(x|\theta_0)}$ 為遞增函數。

$\Leftrightarrow \dfrac{\partial}{\partial x}\left\{\dfrac{f(x|\theta_1)}{f(x|\theta_0)}\right\} = \dfrac{f_x(x|\theta_1)f(x|\theta_0) - f(x|\theta_1)f_x(x|\theta_0)}{\{f(x|\theta_0)\}^2} \geqq 0$

$\Leftrightarrow f_x(x|\theta_1)f(x|\theta_0) \geqq f(x|\theta_1)f_x(x|\theta_0)$

$\Leftrightarrow \dfrac{f_x(x|\theta_1)}{f(x|\theta_1)} \geqq \dfrac{f_x(x|\theta_0)}{f(x|\theta_0)}$ （註 $f_x = \dfrac{\partial f}{\partial x}$, $f_\theta = \dfrac{\partial f}{\partial \theta}$, $f_{x\theta} = \dfrac{\partial^2 f}{\partial x \partial \theta}$）

$\Leftrightarrow \dfrac{f_x(x|\theta)}{f(x|\theta)}$ 為 θ 之遞增函數。（x 固定的情況下）

$\dfrac{\partial}{\partial \theta}\dfrac{f_x(x|\theta)}{f(x|\theta)} = \dfrac{f_{x\theta}(x|\theta)f(x|\theta) - f_x(x|\theta)f_\theta(x|\theta)}{\{f(x|\theta)\}^2} \geqq 0$

$\Leftrightarrow f_{x\theta}(x|\theta)f(x|\theta) \geqq f_x(x|\theta)f_\theta(x|\theta)$ …… $(*)$ 以下證明 ①$\Leftrightarrow(*)$

以及 ②$\Leftrightarrow(*)$

① 若 $\dfrac{\partial^2}{\partial\theta\partial x}\log f(x|\theta) \geqq 0$ $(\forall x, \theta)$

$\Leftrightarrow \dfrac{\partial}{\partial \theta}\dfrac{f_x(x|\theta)}{f(x|\theta)} = \dfrac{f_{x\theta}(x|\theta)f(x|\theta) - f_x(x|\theta)f_\theta(x|\theta)}{\{f(x|\theta)\}^2} \geqq 0$

$\Leftrightarrow f_{x\theta}(x|\theta)f(x|\theta) \geqq f_x(x|\theta)f_\theta(x|\theta)$

$\Leftrightarrow (*)$ ∴ 證明完成

② 題目應該有錯：

$$f(x|\theta)\,\frac{\partial}{\partial\theta}\left(\frac{\partial}{\partial x}\log f(x|\theta)\right) \geq \frac{\partial f(x|\theta)}{\partial\theta}\,\frac{\partial f(x|\theta)}{\partial x}$$

$$\Leftrightarrow\ f(x|\theta)\left\{\frac{f_{x\theta}(x|\theta)f(x|\theta)-f_x(x|\theta)f_\theta(x|\theta)}{(f(x|\theta))^2}\right\} \geq f_\theta(x|\theta)f_x(x|\theta)$$

$$\Leftrightarrow\ f_{x\theta}(x|\theta)f(x|\theta)-f_x(x|\theta)f_\theta(x|\theta) \geq f_\theta(x|\theta)f_x(x|\theta)f(x|\theta)$$

這跟 ⊛ 並非等價。

題目應改為 $f(x|\theta)\cdot\underbrace{\frac{\partial^2}{\partial\theta\partial x}f(x|\theta)}_{\text{(去掉 log)}} \geq \frac{\partial f(x|\theta)}{\partial\theta}\cdot\frac{\partial f(x|\theta)}{\partial x}$

$$\Leftrightarrow\ f_{x\theta}(x|\theta)f(x|\theta) \geq f_x(x|\theta)f_\theta(x|\theta)$$

$\Leftrightarrow$ ⊛

Date　　　　Page

(2) 先求之元機率密度函數

$Pr(Z\leq x)=Pr(-\sqrt{x}\leq Z\leq\sqrt{x})=Pr(-\sqrt{x}-\sqrt{\theta}\leq Z-\sqrt{\theta}\leq\sqrt{x}-\sqrt{\theta})$

$=\Phi(\sqrt{x}-\sqrt{\theta})-\Phi(-\sqrt{x}-\sqrt{\theta})$

$\frac{d}{dx}Pr(Z^2\leq x)=\frac{1}{2\sqrt{x}}\phi(\sqrt{x}-\sqrt{\theta})+\frac{1}{2\sqrt{x}}\phi(-\sqrt{x}-\sqrt{\theta})$

$=\frac{1}{2\sqrt{x}\sqrt{2\pi}}\exp\left(\frac{-x}{2}-\frac{\theta}{2}\right)\left\{\exp(\sqrt{\theta x})+\exp(-\sqrt{\theta x})\right\}\quad(x>0)$

$\therefore f(x|\theta)=\frac{1}{2\sqrt{x}\sqrt{2\pi}}\exp\left(\frac{-x}{2}-\frac{\theta}{2}\right)\left\{\exp(\sqrt{\theta x})+\exp(-\sqrt{\theta x})\right\}\quad(x>0)$

$\frac{f(x|\theta_1)}{f(x|\theta_0)}=\exp\left(-\frac{\theta_1}{2}+\frac{\theta_0}{2}\right)\cdot\frac{\exp(\sqrt{\theta_1 x})+\exp(-\sqrt{\theta_1 x})}{\exp(\sqrt{\theta_0 x})+\exp(-\sqrt{\theta_0 x})}\quad(x>0)$

$a_1\overset{def}{=}\sqrt{\theta_1}\quad a_0\overset{def}{=}\sqrt{\theta_0}\quad y\overset{def}{=}\sqrt{x}$

$\Lambda=\frac{f(x|\theta_1)}{f(x|\theta_0)}=\exp\left(-\frac{\theta_1}{2}+\frac{\theta_0}{2}\right)\cdot\frac{\cosh(a_1y)}{\cosh(a_0y)}$

我們先證明 Λ 與 y 之間有遞增或遞減的關係

$\lambda(y)\overset{def}{=}\frac{\cosh(a_1y)}{\cosh(a_0y)}$

$\lambda'(y)=\frac{a_1\sinh(a_1y)\cosh(a_0y)-a_0\cosh(a_1y)\sinh(a_0y)}{(\cosh(a_0y))^2}$

$$g(y) \overset{def}{=} a_1 \sinh(a_1 y)\cosh(a_0 y) - a_0 \cosh(a_1 y)\sinh(a_0 y) \quad (y>0)$$

$$g'(y) = a_1^2 \cosh(a_1 y)\cosh(a_0 y) + a_0 a_1 \sinh(a_1 y)\sinh(a_0 y)$$

$$- a_0^2 \cosh(a_1 y)\cosh(a_0 y) - a_0 a_1 \sinh(a_1 y)\sinh(a_0 y)$$

$$= (a_1^2 - a_0^2)\cosh(a_1 y)\cosh(a_0 y) > 0 \quad (y>0) \quad (\text{if } \theta_1 > \theta_0)$$

$g(y)$ 為遞增函數. $g(0)=0$. $\therefore g(y)>0$

$\Rightarrow \lambda'(y) > 0 \Rightarrow \lambda(y)$ 為遞增函數

$\therefore$ y 增加 $\Leftrightarrow$ λ 增加

$\Updownarrow$

Λ 增加

$\therefore$ 證明完成.

Date Page

$\boxed{15}$

(1) 考慮對於 $H_0: \theta=\theta_0$ vs $H_1: \theta=\theta_1$ 之最強力檢定。$\overset{\text{假設}}{(\theta_0<\theta_1)}$

利用 Neyman-Pearson's lemma:

$C=\left\{x \,\middle|\, \frac{f(x|\theta_1)}{f(x|\theta_0)}>k\right\}$ 為 MP 檢定的棄卻域。

由於 $f(x)$ 具有單調概似比，

x 增加 $\Leftrightarrow$ $\frac{f(x|\theta_1)}{f(x|\theta_0)}$ 增加

$\therefore C=\{x \mid x>x_0\}$

$\varphi_c(x)=\begin{cases}1 & x\in C\\ 0 & x\notin C\end{cases}$ 為最強力檢定。

但 $E[\varphi_c(x)\mid\theta=\theta_0]=\alpha$ （顯著水準）

另外，考慮隨機化檢定。

$\varphi_c^*(x)=\begin{cases}1 & (\text{機率 }\alpha)\\ 0 & (\text{〃 } 1-\alpha)\end{cases}$

$\varphi_c^*(x)$ 亦為顯著水準 α 之檢定

φ_c 之檢定力 $\geq$ φ_c^* 之檢定力

$\therefore E[\varphi_c(X)|\theta_1] \geq E[\varphi_c^*(X)|\theta_1] = \alpha = E[\varphi_c(X)|\theta_0]$

故此 $E[\varphi_c(X)|\theta_1] \geq E[\varphi_c(X)|\theta_0]$

$\Leftrightarrow \int_{x>x_0} f(x|\theta_1)\,dx \geq \int_{x>x_0} f(x|\theta_0)\,dx$

$\Leftrightarrow 1-F(x_0|\theta_1) \geq 1-F(x_0|\theta_0)$

$\Leftrightarrow F(x_0|\theta_0) \geq F(x_0|\theta_1)$

$\therefore$ stochastically increasing

(根據 x_0, 我們可以選擇不同的 α)

Red Elephant

(2) 證明存在 stochastically increasing

卻不具 MLR 的分佈.

考慮以下分佈: $(\theta_0 < \theta_1)$

X	0	1	2
$Pr(X\|\theta_0)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$Pr(X\|\theta_1)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

$$Pr(X \leq 0 \mid \theta_0) = \frac{1}{3} \geq Pr(X \leq 0 \mid \theta_1) = \frac{1}{3}$$

$$Pr(X \leq 1 \mid \theta_0) = \frac{2}{3} \geq Pr(X \leq 1 \mid \theta_1) = \frac{1}{2}$$

$$Pr(X \leq 2 \mid \theta_0) = 1 \geq Pr(X \leq 2 \mid \theta_1) = 1$$

由此可知, X is stochastically increasing.

$$\frac{Pr(X=0 \mid \theta_1)}{Pr(X=0 \mid \theta_0)} = 1 \underset{>}{,} \frac{Pr(X=1 \mid \theta_1)}{Pr(X=1 \mid \theta_0)} = 0.5 \underset{<}{,} \frac{Pr(X=2 \mid \theta_1)}{Pr(X=2 \mid \theta_0)} = 1.5$$

由此可知, X 並不具 MLR.

∴ 證明完成.