Homework 7 Solution

1.
(Problem 3.3.15)
Jeffrey's "Prior". A density proportional to $\sqrt{I_p(\theta)}$ is called Jeffery's prior. It is often improper. Show that in the $N(\theta, \sigma_0^2), N(\mu_0, \theta)$ and $B(n, \theta)$ cases, Jeffery's priors are proportional to $1, \theta^{-1}$, and $\theta^{-\frac{1}{2}}(1-\theta)^{-\frac{1}{2}}$, respectively. Give the Bayes rules for squared error in these three cases.
Solution:
(i) For $N(\theta, \sigma_0^2)$,
$L(\theta) = f_\theta(x) = (2\pi\sigma_0^2)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma_0^2}(x-\theta)^2\}$,
Then,
$\log L(\theta) = -\frac{1}{2}\log(2\pi\sigma_0^2) - \frac{1}{2\sigma_0^2}(x-\theta)^2$,
$\Rightarrow \frac{\partial \log L(\theta)}{\partial\theta} = \frac{1}{\sigma_0^2}(x-\theta)$
$\Rightarrow I_p(\theta) = E\left(\frac{\partial \log L(\theta)}{\partial\theta}\right)^2 = \frac{1}{\sigma_0^4}E(x-\theta)^2 = \frac{1}{\sigma_0^2}$
So, the Jeffery's prior density of θ is in the form of
$g(\theta) \propto \sqrt{I_p(\theta)} = \frac{1}{\sigma_0}$, i.e. $g(\theta) \propto 1$;
As $\pi(\theta|x) \tilde{} N(\bar{x}, \sigma_0^2)$, the Bayes rule for squared loss is $\bar{x}$.
(ii) For $N(\mu_0, \theta)$,
$L(\theta) = f_\theta(x) = (2\pi\theta)^{-\frac{1}{2}} \exp\{-\frac{1}{2\theta}(x-\mu_0)^2\}$,
Then,
$\log L(\theta) = -\frac{1}{2}\log(2\pi\theta) - \frac{1}{2\theta}(x-\mu_0)^2$,
$\Rightarrow \frac{\partial \log L(\theta)}{\partial\theta} = -\frac{1}{2\theta} + \frac{1}{2\theta^2}(x-\mu_0)^2$
$\Rightarrow \frac{\partial^2 \log L(\theta)}{\partial\theta^2} = \frac{1}{2\theta^2} - \frac{1}{\theta^3}(x-\mu_0)^2$
$\Rightarrow I_p(\theta) = -E\left(\frac{\partial^2 \log L(\theta)}{\partial\theta^2}\right) = -\frac{1}{2\theta^2} + \frac{1}{\theta^2} = \frac{1}{\theta^2}$
So the Jeffery's prior density of θ is in the form of
$g(\theta) \propto \sqrt{I_p(\theta)} = \frac{1}{\theta}$;
As $\pi(\theta|x) \propto \theta^{-\left(\frac{n}{2}+1\right)} \exp\left\{-\frac{\sum_{i=1}^n (x_i-\mu_0)^2}{2\theta}\right\}$, which indicates that $\theta|x$ follows an inverse gamma distribution. So the Bayes rule for the squared loss is $\frac{\sum_{i=1}^n (x_i-\mu_0)^2}{n}$.
(iii) For $B(n, \theta)$,
$L(\theta) = f_\theta(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$,
Then,
$\log L(\theta) = \log\binom{n}{x} + x\log\theta + (n-x)\log(1-\theta)$,
$\Rightarrow \frac{\partial \log L(\theta)}{\partial\theta} = \frac{x}{\theta} - \frac{n-x}{1-\theta}$
$\Rightarrow \frac{\partial^2 \log L(\theta)}{\partial\theta^2} = -\frac{x}{\theta^2} - \frac{n-x}{(1-\theta)^2}$
$\Rightarrow I_p(\theta) = -E\left(\frac{\partial^2 \log L(\theta)}{\partial\theta^2}\right) = \frac{E(x)}{\theta^2} + \frac{E(n-x)}{(1-\theta)^2} = \frac{n}{\theta} + \frac{n}{1-\theta} = \frac{n}{\theta(1-\theta)}$
So the Jeffery's prior density of θ is in the form of
$g(\theta) \propto \sqrt{I_p(\theta)} = \frac{\sqrt{n}}{\sqrt{\theta(1-\theta)}}$,
i.e. $g(\theta) \propto \theta^{-\frac{1}{2}}(1-\theta)^{-\frac{1}{2}}$.
As $\pi(\theta|x) \tilde{} Beta\left(\sum_{i=1}^m x_i + \frac{1}{2}, mn - \sum_{i=1}^m x_i + \frac{1}{2}\right)$, the Bayes rule is $\frac{\sum_{i=1}^m x_i + \frac{1}{2}}{mn+1}$.

(Problem 3.4.11)
Suppose $Y_1, ..., Y_n$ are independent Poisson random variables with $E(Y_i) = \mu_i$ where $\mu_i = \exp\{\alpha + \beta z_i\}$ depends on the levels z_i of a covariate; $\alpha, \beta \in R$. For instance, z_i could be the level of a drug given to the ith patient with an infectious disease and Y_i could denote the number of infectious agents in a given unit of blood from the ith patient 24 hours after the drug was administered.
(a) Write the model for $Y_1, ..., Y_n$ in two-parameter canonical exponential form and give the sufficient statistic.
(b) Let $\theta = (\alpha, \beta)^T$. Compute $I(\theta)$ for the model in (a) and then find the lower bound on the variances of unbiased estimators $\hat{\alpha}$ and $\hat{\beta}$ of α and β.
(c) Suppose that $z_i = \log[i/(n+1)], i = 1, ..., n$. Find $\lim n^{-1} I(\theta)$ as $n \to \infty$, and give the limit of n times the lower bound on the variances of $\hat{\alpha}$ and $\hat{\beta}$.
Hint: Use the integral approximation to sums.
Solution:

(a) As $Y_i \tilde{} Poisson(\mu_i)$, with $\mu_i = \exp\{\alpha + \beta z_i\}$,

$$f_{\alpha,\beta}(y_1, ..., y_n) = e^{-\sum_{i=1}^n \mu_i} \prod_{i=1}^n \frac{\mu_i^{y_i}}{y_i!}$$

$$= e^{-\sum_{i=1}^n \exp(\alpha+\beta z_i)} \frac{\exp \sum_{i=1}^n (\alpha+\beta z_i) y_i}{\prod_{i=1}^n y_i!} = \exp\left\{-\sum_{i=1}^n \exp(\alpha + \beta z_i) + \sum_{i=1}^n (\alpha + \beta z_i) y_i - \sum_{i=1}^n \log y_i!\right\}$$

$$= \exp\left\{\alpha \sum_i y_i + \beta \sum_i y_i z_i - \sum_{i=1}^n \exp(\alpha + \beta z_i) - \sum_{i=1}^n \log y_i!\right\}$$

$$= \exp\left\{\eta_1 T_1(\underline{y}) + \eta_2 T_2(\underline{y}) - \xi(\eta_1, \eta_2)\right\} h(\underline{y})$$

Where,

$\eta_1 = \alpha, \eta_2 = \beta,$
$T_1(\underline{y}) = \sum_i y_i, T_2(\underline{y}) = \sum_i y_i z_i,$
$\xi(\eta_1, \eta_2) = \sum_{i=1}^n \exp(\alpha + \beta z_i),$
$h(\underline{y}) = \exp\{-\sum_{i=1}^n \log y_i!\} = \frac{1}{\prod_{i=1}^n y_i!};$

Hence, the sufficient statistics are $T_1(\underline{y}) = \sum_i y_i, T_2(\underline{y}) = \sum_i y_i z_i$.
(b) $L(\alpha, \beta)$

$$= \exp\left\{\alpha \sum_i y_i + \beta \sum_i y_i z_i - \sum_{i=1}^n \exp(\alpha + \beta z_i) - \sum_{i=1}^n \log y_i!\right\}$$

$$\Rightarrow \log L(\alpha, \beta) = \alpha \sum_i y_i + \beta \sum_i y_i z_i - \sum_{i=1}^n \exp(\alpha + \beta z_i) - \sum_{i=1}^n \log y_i!$$

$$\Rightarrow I_{11}(\theta) = -E\left(\frac{\partial^2 \log L(\alpha,\beta)}{\partial \alpha^2}\right) = \sum_{i=1}^n \exp(\alpha + \beta z_i)$$

$$I_{12}(\theta) = -E\left(\frac{\partial^2 \log L(\alpha,\beta)}{\partial \alpha \partial \beta}\right) = \sum_{i=1}^n z_i \exp(\alpha + \beta z_i)$$

$$I_{22}(\theta) = -E\left(\frac{\partial^2 \log L(\alpha,\beta)}{\partial \beta^2}\right) = \sum_{i=1}^n z_i^2 \exp(\alpha + \beta z_i)$$

$$\Rightarrow I(\theta) = \begin{pmatrix} \sum_{i=1}^n \exp(\alpha + \beta z_i) & \sum_{i=1}^n z_i \exp(\alpha + \beta z_i) \\ \sum_{i=1}^n z_i \exp(\alpha + \beta z_i) & \sum_{i=1}^n z_i^2 \exp(\alpha + \beta z_i) \end{pmatrix};$$

So, the lower bound of the variance for $\hat{\alpha}$ is $\frac{\sum_{i=1}^n z_i^2 \exp(\alpha+\beta z_i)}{\det(I(\theta))}$ and for $\hat{\beta}$ is $\frac{\sum_{i=1}^n \exp(\alpha+\beta z_i)}{\det(I(\theta))}$, where $\det(I(\theta)) = \left(\sum_{i=1}^n \exp(\alpha + \beta z_i)\right)\left(\sum_{i=1}^n z_i^2 \exp(\alpha + \beta z_i)\right) - \left(\sum_{i=1}^n z_i \exp(\alpha + \beta z_i)\right)^2;$
(c) $n^{-1} I(\theta)$

$$= \begin{pmatrix} \sum_i \frac{1}{n} \exp\left(\alpha + \beta \log \frac{i}{n+1}\right) & \sum_i \frac{1}{n} \log \frac{i}{n+1} \exp\left(\alpha + \beta \log \frac{i}{n+1}\right) \\ \sum_i \frac{1}{n} \log \frac{i}{n+1} \exp\left(\alpha + \beta \log \frac{i}{n+1}\right) & \sum_i \frac{1}{n} \left(\log \frac{i}{n+1}\right)^2 \exp\left(\alpha + \beta \log \frac{i}{n+1}\right) \end{pmatrix}$$

As $\lim_{n\to\infty} \sum_i \frac{1}{n} \exp\left(\alpha + \beta \log \frac{i}{n+1}\right) = \exp \alpha (\int_0^1 x^\beta dx) = \frac{e^\alpha}{\beta+1}$,

$\lim_{n\to\infty} \sum_i \frac{1}{n} \log \frac{i}{n+1} \exp\left(\alpha + \beta \log \frac{i}{n+1}\right)$

$= \exp\alpha\left(\int_0^1 x^\beta \log x dx\right) = -\frac{e^\alpha}{(\beta+1)^2}$,

$\lim_{n\to\infty} \sum_i \frac{1}{n}\left(\log\frac{i}{n+1}\right)^2 \exp\left(\alpha+\beta\log\frac{i}{n+1}\right)$

$= \exp\alpha\left(\int_0^1 x^\beta (\log x)^2 dx\right) = \frac{2e^\alpha}{(\beta+1)^3}$,

$\Rightarrow \lim_{n\to\infty} n^{-1} I(\theta) = \begin{pmatrix} \frac{e^\alpha}{\beta+1} & -\frac{e^\alpha}{(\beta+1)^2} \\ -\frac{e^\alpha}{(\beta+1)^2} & \frac{2e^\alpha}{(\beta+1)^3} \end{pmatrix}$;

So the limites of n times the lower bound for $\hat{\alpha}$ is $(\beta+1)e^{-\alpha}$, and for $\hat{\beta}$ is $(\beta+1)^3 e^{-\alpha}/2$.

(Problem 3.4.12)

Let $X_1, ..., X_n$ be a sample from the beta, $B(\theta, 1)$, distribution.

(a) Find the MLE of $\frac{1}{\theta}$. Is it unbiased? Does it achieve the information inequality lower bound?

(b) Show that $\bar{X}$ is an unbiased estimate of $\frac{\theta}{\theta+1}$. Does $\bar{X}$ achieve the information inequality lower bound?

Solution:

(a) $L(\theta) = f_\theta(\underline{X}) = \prod_{i=1}^n \frac{1}{B(\theta,1)} x_{i^{\theta-1}} = \theta^n\left(\prod_{i=1}^n x_i\right)^{\theta-1}$

$\Rightarrow \log L(\theta) = n\log\theta + (\theta-1)\sum_{i=1}^n \log x_i$

By setting $\frac{\partial \log L(\theta)}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^n \log x_i = 0$

$\Rightarrow \hat{\theta} = -\frac{n}{\sum_{i=1}^n \log x_i}$, i.e. $\widehat{\frac{1}{\theta}} = -\frac{1}{n}\sum_{i=1}^n \log x_i$;

Since Beta distribution belongs to the exponential family

$f(x|\theta) = \exp\{(\theta-1)\log x + \log\theta\}$,

$\Rightarrow E\log x = -\frac{1}{\theta}$

Then, $\hat{\theta}_{MLE}^{-1}$ is unbiased and it achieves the information inequality lower bound.

(b) $Var(\log x) = I(\theta) = \frac{1}{\theta^2}$, and $E\bar{x} = \frac{\theta}{\theta+1}$,

By the property of Beta distribution,

$Var\left(\frac{\sum_i X_i}{n}\right) = \frac{1}{n}\frac{\theta}{(\theta+1)^2(\theta+2)} > \frac{\left[\frac{d}{d\theta}\left(\frac{\theta}{\theta+1}\right)\right]^2}{nI(\theta)} = \frac{\theta^2}{n(1+\theta)^4}$,

Then, $\bar{x}$ does not achieve the information inequality lower bound.

(Problem 3.4.22)

Regularity Conditions are Needed for the Information Inequality. Let $X\tilde{\ }U(0,\theta)$ be the uniform distribution on $(0,\theta)$. Note that $\log p(x,\theta)$ is differentiable for all $\theta > x$, that is, with probability 1 for each θ, and we can thus define moments of $\frac{\partial \log p(x,\theta)}{\partial\theta}$. Show that, however,

(i) $E\left(\frac{\partial}{\partial\theta}\log p(x,\theta)\right) = -\frac{1}{\theta} \neq 0$;

(ii) $Var\left(\frac{\partial}{\partial\theta}\log p(x,\theta)\right) = 0$ and the information bound is infinite. Yet show

(iii) $2X$ is unbiased for θ and has finite variance.

Solution:

(i) $P(x|\theta) = \frac{1}{\theta} I(0 < x < \theta)$

$\Rightarrow \frac{\partial \log P(x|\theta)}{\partial\theta} = -\frac{1}{\theta}$

$\Rightarrow E\frac{\partial \log P(x|\theta)}{\partial\theta} = -\frac{1}{\theta} \neq 0$;

(ii) $Var\left(\frac{\partial \log P(x|\theta)}{\partial\theta}\right) = Var\left(-\frac{1}{\theta}\right) = 0$
So, the information bound is infinite;
(iii) $E2X = 2EX = \theta$
$\Rightarrow 2X$ is unbiased for θ.
$Var2X = 4VarX = \frac{\theta^2}{3} < \infty$.

2. Suppose that $X\tilde{\ }N(\mu, \sigma^2)$.
a) Consider the two-dimensianl parameter, $\theta = (\mu, \sigma)$. Find the Fisher information matrix $I_X(\theta_0)$.
b) Then consider the reparameterization in exponential family form with
$\eta = \binom{\eta_1}{\eta_2} = \binom{\frac{\mu}{\sigma^2}}{-\frac{1}{2\sigma^2}} = \binom{\frac{\theta_1}{\theta_2^2}}{-\frac{1}{2\theta_2^2}}$
What is $I_X(\eta_0)$?
c) If

$$g(\theta) = \binom{\frac{\theta_1}{\theta_2^2}}{-\frac{1}{2\theta_2^2}}$$

and $\theta_0 = g^{-1}(\eta_0)$, can one simply plug θ_0 into matrix from a) to get the matrix from b)? The complete story hinted at here is told in Problem 3.4.3 of B&D.
Solution:
a) For $X\tilde{\ }N(\mu, \sigma^2), \theta = (\mu, \sigma)$,

$$L(\theta) = f_\theta(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\},$$

$$\Rightarrow \log L(\theta) = -\frac{1}{2}\log(2\pi) - \log\sigma - \frac{1}{2\sigma^2}(x-\mu)^2,$$

$$\Rightarrow I_{11}(\theta_0) = -E\left(\frac{\partial^2 \log L(\theta)}{\partial\mu^2}|_{\theta=\theta_0}\right) = \frac{1}{\sigma_0^2},$$

$$I_{12}(\theta_0) = -E\left(\frac{\partial^2 \log L(\theta)}{\partial\mu\partial\sigma}|_{\theta=\theta_0}\right) = \frac{2}{\sigma_0^3}E_{\theta_0}(x-\mu_0) = 0,$$

$$I_{22}(\theta_0) = -E\left(\frac{\partial^2 \log L(\theta)}{\partial\sigma^2}|_{\theta=\theta_0}\right)$$

$$= -\frac{1}{\sigma_0^2} + \frac{3}{\sigma_0^4}E_{\theta_0}(x-\mu_0)^2 = -\frac{1}{\sigma_0^2} + \frac{3}{\sigma_0^2} = \frac{2}{\sigma_0^2};$$

So the Fisher information matrix

$$I_X(\theta_0) = \begin{pmatrix} I_{11}(\theta_0) & I_{12}(\theta_0) \\ I_{12}(\theta_0) & I_{22}(\theta_0) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_0^2} & 0 \\ 0 & \frac{2}{\sigma_0^2} \end{pmatrix}.$$

b) After reparameterization in exponential family form with

$$\eta = \binom{\eta_1}{\eta_2} = \binom{\frac{\mu}{\sigma^2}}{-\frac{1}{2\sigma^2}} = \binom{\frac{\theta_1}{\theta_2^2}}{-\frac{1}{2\theta_2^2}},$$

$$L(\eta) = f_\eta(x) = \left(-2\pi\frac{1}{2\eta_2}\right)^{-\frac{1}{2}} \exp\{\eta_2\left(x + \frac{\eta_1}{2\eta_2}\right)^2\},$$

$$\Rightarrow \log L(\eta) = -\frac{1}{2}\log\left(-\frac{\pi}{\eta_2}\right) + \eta_2\left(x + \frac{\eta_1}{2\eta_2}\right)^2,$$

$$\Rightarrow I_{11}(\eta_0) = -E\left(\frac{\partial^2 \log L(\eta)}{\partial\eta_1^2}|_{\eta=\eta_0}\right) = -\frac{1}{2\eta_2},$$

$$I_{12}(\eta_0) = -E\left(\frac{\partial^2 \log L(\eta)}{\partial\eta_1\partial\eta_2}|_{\eta=\eta_0}\right) = \frac{\eta_1}{2\eta_2^2},$$

$$I_{22}(\eta_0) = -E\left(\frac{\partial^2 \log L(\eta)}{\partial\eta_2^2}|_{\eta=\eta_0}\right) = \frac{1}{2\eta_2^2} - \frac{\eta_1^2}{2\eta_2^3}$$

So the Fisher information matrix

$$I_X(\eta_0) = \begin{pmatrix} I_{11}(\eta_0) & I_{12}(\eta_0) \\ I_{12}(\eta_0) & I_{22}(\eta_0) \end{pmatrix} = \begin{pmatrix} -\frac{1}{2\eta_2} & \frac{\eta_1}{2\eta_2^2} \\ \frac{\eta_1}{2\eta_2^2} & \frac{1}{2\eta_2^2} - \frac{\eta_1^2}{2\eta_2^3} \end{pmatrix}.$$

c) From the two information matrix above, we see that it is impossible to simply plug θ_0 into matrix from a) to get the matrix from b).

3. What is the "Jeffery's prior" for a model with $X\tilde{\ }Poisson\,(\lambda)$? For this "prior", what is the posterior distribution of $\lambda|X$?
Solution:
For $X\tilde{\ }Poisson\,(\lambda)$,
$L(\lambda) = f_\lambda(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \lambda > 0, x \in Z^+,$
Then,
$\log L(\lambda) = -\lambda + x\log\lambda - \log(x!)\,,$
$\Rightarrow \frac{\partial \log L(\lambda)}{\partial\lambda} = -1 + \frac{x}{\lambda},$
$\Rightarrow I_X(\lambda) = E\left(\frac{\partial \log L(\lambda)}{\partial\lambda}\right)^2 = \frac{1}{\lambda^2}E(x-\lambda)^2 = \frac{1}{\lambda}.$
So the "Jeffery's prior" for λ is in the form of
$g(\lambda) \propto \sqrt{I_X(\lambda)} = \frac{1}{\sqrt{\lambda}}.$
For this "prior", the posterior distribution for $\lambda|X$ is like
$g(\lambda|x) \propto L(\lambda)\,g(\lambda) = e^{-\lambda}\frac{\lambda^x}{x!}\frac{1}{\sqrt{\lambda}} = e^{-\lambda}\frac{\lambda^{x-\frac{1}{2}}}{x!},$
Where, $e^{-\lambda}\lambda^{x-\frac{1}{2}}$ is the kernel of $Gamma\left(x+\frac{1}{2},1\right)$.
Hence, $\lambda|X\tilde{\ }Gamma\left(x+\frac{1}{2},1\right)$.

4. (Optional only, highly recommended but not required)
Suppose that X is a discrete random variable taking values in some finite set $\aleph$ and that $f(x|\theta) > 0$ for all x for $\theta \in \Theta$ (some open interval in $\Re$). Suppose further for each $x \in \aleph$ that $f(x|\theta)$ is differentiable in θ at θ_0. Show that for any statistic $T(X)$,
$$I_{T(X)}(\theta_0) \leq I_X(\theta_0)\,.$$
Solution:
$L(\theta) = f(x|\theta)\,,$

$I_X(\theta) = E(\frac{\partial^2 \log L(\theta)}{\partial\theta^2}) = E\left\{E\left(\frac{\partial^2 \log L(\theta)}{\partial\theta^2}|T(x)=t\right)\right\}$
$\geq E\left\{E\left(\frac{\partial \log L(\theta)}{\partial\theta}|T(x)=t\right)\right\}^2; \quad \text{(by Jensen's)} \qquad (*)$
As $P(X=x|T(x)=t) = \frac{P(X=x)}{\sum_{x\in\Im}P(X=x)}I(x\in\Im)\,, \quad (1)$
where, $\Im = \{x|T(x)=t\}\,,$
then,
$P(T(x)=t) = \sum_{x\in\Im}P(X=x)$
$\Rightarrow \log L(\theta|T(X)) = \log\sum_{x\in\Im}P(X=x)$
$\Rightarrow \frac{\partial \log L(\theta|T(X))}{\partial\theta} = \frac{\sum_{x\in\Im}\frac{d}{d\theta}P(X=x)}{\sum_{x\in\Im}P(X=x)}; \qquad (2)$
By (1) & (2),
$E\left(\frac{\partial \log L(\theta)}{\partial\theta}|T(X)=t\right) = \frac{\sum_{x\in\Im}\frac{\partial \log L(\theta)}{\partial\theta}P(X=x)}{\sum_{x\in\Im}P(X=x)}$
$= \frac{\sum_{x\in\Im}\frac{d}{d\theta}P(X=x)}{\sum_{x\in\Im}P(X=x)} = \frac{\partial \log L(\theta|T(X))}{\partial\theta},$

And by $(*)$,
$I_X(\theta) \geq E\left\{E\left(\frac{\partial \log L(\theta)}{\partial\theta}|T(x)=t\right)\right\}^2$

$$= E\left(\frac{\partial \log L(\theta|T(X))}{\partial\theta}\right)^2 = I_{T(X)}(\theta).$$

As $f(x|\theta)$ is differentiable in θ at θ_0, then we can finish the proof by giving $\theta = \theta_0$.

5. (Optional only, recommended but not required)
(Problem 3.4.3)
Equivariance. Let $X\tilde{\ }p(x,\theta)$ with $\theta \in \Theta \subset R$, suppose that assumptions I and II hold and that h is a monotone increasing diffrentiable function from Θ onto $h(\Theta)$. Reparametrize the model by setting $\eta = h(\theta)$ and let $q(x,\eta) = p(x, h^{-1}(\eta))$ denote the model in the new parametrization.
(a) Show that if $I_p(\theta)$ and $I_q(\theta)$ denote the Fisher information in the two arametrizations, then
$I_q(\eta) = I_p(h^{-1}(\eta)) / \left[h'(h^{-1}(\eta))\right]^2.$
That is, Fisher information is not equivariant under increasing transformations of the parameter.
(b) *Equivariance of the Fisher Information Bound.* Let $B_p(\theta)$ and $B_q(\eta)$ denote the information inequality lower bound $(\psi')^2/I$ as in (3.4.12) for the two parametrizations $p(x,\theta)$ and $q(x,\eta)$. Show that $B_q(\eta) = B_p(h^{-1}(\eta))$; that is, the Fisher information lower bound is equivariant.
Solution:

(a) $I_p(\theta) = E\left(\frac{d\log p(x,\theta)}{d\theta}\right)^2,$
$\theta = h^{-1}(\eta)$ is monotone, and $q(x,\eta) = p(x,h^{-1}(\eta))$.

$$I_q(\eta) = E\left(\frac{d\log p\left(x,h^{-1}(\eta)\right)}{d\eta}\right)$$
$$= E\left(\frac{d\log p\left(x,h^{-1}(\eta)\right)}{dh^{-1}(\eta)}\frac{dh^{-1}(\eta)}{d\eta}\right)^2$$
$$= \frac{I_p\left(h^{-1}(\eta)\right)}{(h'(h^{-1}(\eta)))^2};$$

(b) $Var_\theta(\delta(x)) \geq \frac{\left(\frac{dE\delta(x)}{d\theta}\right)^2}{I_p(\theta)} = B_p(\theta)$
$Var_\eta(\delta(x)) \geq \frac{\left(\frac{dE\delta(x)}{d\eta}\right)^2}{I_q(\eta)} = B_q(\eta)$

$$\begin{aligned}\frac{d}{d\eta}E\delta(x) &= \frac{d}{d\eta}\int\delta(x)\,p(x,h^{-1}(\eta))\,dx\\ &= \int\delta(x)\frac{d}{d\eta}p(x,h^{-1}(\eta))\,dx\\ &= \int\delta(x)\frac{d}{dh^{-1}(\eta)}p(x,h^{-1}(\eta))\frac{dh^{-1}(\eta)}{d\eta}dx\\ &= \frac{1}{h'(h^{-1}(\eta))}\frac{dE\delta(x)}{d\theta}\Big|_{\theta=h^{-1}(\eta)}\end{aligned}$$
$$\Rightarrow \frac{\left(\frac{dE\delta(x)}{d\eta}\right)^2}{I_q(\eta)} = \frac{1}{(h'(h^{-1}(\eta)))^2}\frac{\left(\frac{dE\delta(x)}{d\theta}|_{\theta=h^{-1}(\eta)}\right)^2}{I_q(\eta)} = \frac{\left(\frac{dE\delta(x)}{d\theta}|_{\theta=h^{-1}(\eta)}\right)^2}{I_p(h^{-1}(\eta))}$$
$\Rightarrow B_q(\eta) = B_p(h^{-1}(\eta))$

(Problem 3.4.20)
Suppose X is distributed according to $\{P_\theta : \theta \in \Theta \subset R\}$ and π is a prior distribution for θ such that $E(\theta^2) < \infty$.

(a) Show that $\delta(X)$ is both an unbiased estimate of θ and the Bayes estimate with respect to quadratic loss, if and only if, $P[\delta(X) = \theta] = 1$.
(b) Deduce that if $P_\theta = N(\theta, \sigma_0^2)$, X is not a Bayes estimate for any prior π.
(c) Explain how it is possible if P_θ is binomial, $B(n, \theta)$, that $\frac{X}{n}$ is a Bayes estimate for θ.
Hint: Given $E(\delta(X)|\theta) = \theta, E(\theta|X) = \delta(X)$ compute $E(\delta(X) - \theta)^2$.
Solution:

(a) " $\Rightarrow$ " :
$P(\delta(x) = \theta) = 1 \Rightarrow \delta(x)$ is both unbiased estimates for θ, and Bayes estimates w.r.t qundratic loss.
" $\Leftarrow$ " :
By definition of unbiased estimates and Bayes estimates w.r.t quadratic loss, we have $E(\delta(x)|\theta) = \theta$ and $E(\theta|x) = \delta(x)$.

$$
\begin{aligned}
& E(\delta(x) - \theta)^2 \\
&= E\left(\delta^2(x) + \theta^2 - 2\delta(x)\theta\right) \\
&= E\left(\delta^2(x) - \delta(x)\theta + \theta^2 - \delta(x)\theta\right) \\
&= E\left\{E\left(\delta^2(x) - \delta(x)\theta\right)|x\right\} + E\left\{E\left(\theta^2 - \delta(x)\theta\right)|x\right\} \\
&= 0
\end{aligned}
$$

$\Rightarrow Var(\delta(x)) = 0$
$\Rightarrow P(\delta(x)) = 1$
(b) As $EX = \theta$, if X is Bayes estimate for some prior π, then
by (a) $\Rightarrow P(X = \theta) = 1$,
which is contradict with $P_0(x) = N(\theta, \sigma_0^2)$
So, X is not a Bayes estimate for any prior π.
(c) If $\theta|X\tilde{\ }Beta(x, n-x)$, then $\frac{X}{n}$ is Bayes estimates for θ.
Consider $\pi \propto \frac{1}{\theta(1-\theta)}$, then

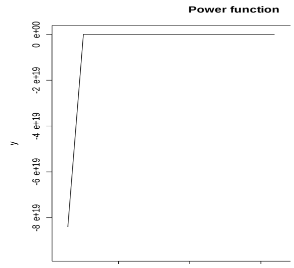
$$f(\theta|x) \propto \theta^x(1-\theta)^{n-x}\theta^{-1}(1-\theta)^{-1} = \theta^{x-1}(1-\theta)^{n-x-1},$$

$\Rightarrow \theta|X\tilde{\ }Beta(x, n-x)$,
So, in the case that prior distribution π is proportional to $\frac{1}{\theta(1-\theta)}$, $\frac{X}{n}$ is Bayes estimates for θ.

6.
(Problem 4.1.1)
Suppose that $X_1, ..., X_n$ are independently and identically distributed according to the uniform distribuion $U(0, \theta)$.
Let $M_n = \max(X_1, ..., X_n)$, and let

$$
\begin{aligned}
\delta_c(X) &= 1 \text{ if } M_n \geq c \\
&= 0 \text{ otherwise.}
\end{aligned}
$$

(a) Compute the power function of δ_c and show that it is a monotone increasing function of θ.
(b) In testing $H : \theta \leq \frac{1}{2}$, what choice of c would make δ_c have size exactly 0.05?
(c) Draw a rough graph of the power function of δ_c specified in (b) when $n = 20$.
(d) How large should n be so that the δ_c specified in (b) has power 0.98 for $\theta = \frac{3}{4}$?
(e) If in a sample of size $n = 20$, $M_n = 0.48$, what is the p-value?
Solution:
(a) The power function of δ_c is



$\delta_c = \beta(\theta) = P_\theta[\delta(X) = 1] = P_\theta[\max(X_1, ..., X_n) \geq c]$
$= \int_c^\theta n\left(\frac{x}{\theta}\right)^{n-1} \frac{1}{\theta} dx = 1 - \left(\frac{c}{\theta}\right)^n;$
As $\frac{\partial\beta(\theta)}{\partial\theta} = n\left(\frac{c}{\theta}\right)^{n-1} \frac{c}{\theta^2} > 0$, then $\beta(\theta)$ is a monotone increasing function of θ.
(b) In testing $H: \theta \leq \frac{1}{2}$, as $\beta(\theta)$ is a monotone increasing function of θ, then
$\delta_c = \sup\{\beta(\theta) : \theta \in \Theta_0\}$
$= \beta\left(\frac{1}{2}\right) = 1 - (2c)^n = 0.05$
$\Rightarrow c = \frac{1}{2}(0.95)^{\frac{1}{n}};$
(c)
(d) $\delta_c = 1 - \left(\frac{c}{\theta}\right)^n = 1 - \left(\frac{\frac{1}{2}(0.95)^{\frac{1}{n}}}{3/4}\right)^n$
$= 1 - 0.95\left(\frac{2}{3}\right)^n \geq 0.98$
$\Rightarrow n \geq \frac{\log 0.02 - \log 0.95}{\log 2 - \log 3} = 9.52$
Take $n = 10$, then the δ_c specified in (b) has power at least 0.98 for $\theta = \frac{3}{4}$.
(e) The p-value is
$p = \alpha(M_n) = \sup\{P_\theta(\max(X_1, ..., X_n) \geq M_n) : \theta \in \Theta_o\}$
$= \sup\{1 - \left(\frac{M_n}{\theta}\right)^n : \theta \leq \frac{1}{2}\}$
$= 1 - \left(\frac{0.48}{1/2}\right)^{20}$
$= 0.558$

(Problem 4.1.3)
Let $X_1, ..., X_n$ be a $\wp(\theta)$ sample.
(a) Use the MLE $\bar{X}$ of θ to construct a level α test for $H: \theta \leq \theta_0$ versus $K: \theta > \theta_0$.
(b) Show that the power function of your test is increasing in θ.
(c) Give an approximate expression for the critical value if n is large and θ is not too close to 0 or ∞. (Use the central limit theorem.)
Solution:

(a) For $X_1, ...X_n$ $i.i.d. Poisson(\theta)$, there exists
$\sum_{i=1}^n X_i \tilde{} Poisson(n\theta)$.
Take the simplest case $n = 2$ as example:
Let $Y = X_1 + X_2$
$X = X_1$
Then,
$X_1 = X$

$X_2 = Y - X$
$\Rightarrow f_{X,Y}(x,y) = f_{X_1,X_2}(x, y-x)\,|J|$
$= e^{-\theta}\frac{\theta^x}{x!}e^{-\theta}\frac{\theta^{y-x}}{(y-x)!}$
$= e^{-2\theta}\frac{\theta^y}{x!(y-x)!}$
$= e^{-2\theta}\frac{\theta^y}{y!}\frac{y!}{x!(y-x)!}$
$\Rightarrow f_Y(y) = \sum_{x=0}^{y} f_{X,Y}(x,y)$
$= e^{-2\theta}\frac{\theta^y}{y!}\sum_{x=0}^{y}\frac{y!}{x!(y-x)!}$
$= e^{-2\theta}\frac{\theta^y}{y!}\sum_{x=0}^{y}\binom{y}{x}$
$= e^{-2\theta}\frac{\theta^y}{y!}$
$\Rightarrow Y\tilde{}\,Poisson\,(2\theta)$
Under the similar argument, we can get
$Y\tilde{}\,Poisson\,(n\theta)$.

Then, by using the MLE $\bar{x}$ of θ, we can construct a level α test for $H_0 : \theta \leq \theta_0\ vs.\ H_1 : \theta > \theta_0$ as

$$\delta(x) = \begin{cases} 1 & \bar{x} \geq \frac{k+1}{n} \\ \gamma & \bar{x} = \frac{k}{n} \\ 0 & \bar{x} < \frac{k}{n} \end{cases}$$

Where, $\gamma = \frac{\alpha - \sum_{x=k+1}^{\infty} P_{n\theta}(x)}{P_{n\theta}(k)}$, and k is determined by
$\sum_{x=k+1}^{\infty} e^{-n\theta}\frac{(n\theta)^x}{x!} \leq \alpha \leq \sum_{x=k}^{\infty} e^{-n\theta}\frac{(n\theta)^x}{x!}$.
(b) $\beta(\theta) = E_\theta\delta(x) = \sum_{x=k+1}^{\infty} e^{-n\theta}\frac{(n\theta)^x}{x!} + \gamma\frac{e^{-n\theta}(n\theta)^k}{k!}$
$= 1 - \sum_{x=1}^{k-1} P_{n\theta}(x) + (\gamma - 1)\,P_{n\theta}(k)$.
If θ is not too small$\left(\theta > \frac{1}{n}\right)$, then
$\frac{d}{d\theta}e^{-n\theta}(n\theta)^x < 0, \ \Rightarrow \beta'(\theta) > 0,$
Hence, $\beta(\theta)$ is increasing in θ.
(c) By the central limit theorem, if n is large and θ is not too close to 0 or ∞,
$\frac{\sqrt{n}(\bar{x}-\theta_0)}{\sqrt{\theta_0}}\tilde{}\,N(0,1)$
$\Rightarrow c = \theta_0 + \sqrt{\frac{\theta_0}{n}}Z_{1-\alpha}.$

(Problem 4.2.1)
Consider Examples 3.3.2 and 4.2.1. You want to buy one of two systems. One has signal-to-noise ratio $v/\sigma_0 = 2$, the other has $v/\sigma_0 = 1$. The first system costs \$$10^6$, the other \$$10^5$. One second of transmission on either system costs \$$10^3$ each. Whichever system you buy during the year, you intend to test the satellite 100 times. If each time you test, you want the number of seconds of response sufficient to ensure that both probabilities of error are ≤ 0.05, which system is cheaper on the basis of a year's operation?
Solution:
Suppose n seconds are needed in each test, and
$X = (x_1, ..., x_n)\,\tilde{}\,N(\mu, \sigma^2)$, with σ^2 known.
Consider testing
$H_0 : \mu = \nu\ vs.\ H_1 : \mu = 0,$
Where ν is a known signal.
Let

$$\delta_c(x) = \begin{cases} 1 & \text{if } \bar{x} \le c \\ 0 & \text{if } \bar{x} > c \end{cases}$$

Then,

$$\begin{cases} P(\delta_c(x) = 1 | H_0) \le 0.05 \\ P(\delta_c(x) = 0 | H_1) \le 0.05 \end{cases}$$

$$\Rightarrow \begin{cases} \frac{\sqrt{n}(x-\nu)}{\sigma_0} \le z_\alpha = -1.645 \\ \frac{\sqrt{n}c}{\sigma_0} \ge z_{1-\alpha} = 1.645 \end{cases}$$

$$\Rightarrow \begin{cases} \sqrt{n} \ge 3.29\frac{\sigma_0}{\nu} \\ \sqrt{n}c \ge 1.645\sigma_0 \end{cases}$$

$$\Rightarrow \begin{cases} n \ge \left(3.29\frac{\sigma_0}{\nu}\right)^2 \\ c \ge \frac{1.645\sigma_0}{\sqrt{n}} \end{cases}$$

So, for the first system, $n \ge \left(\frac{3.29}{2}\right)^2 = 2.7$, take $n = 3$;
and for the seconde system, $n \ge \left(\frac{3.29}{1}\right)^2 = 10.8$, take $n = 11$.
Then, the cost for system 1 is $10^6 + 3 \times 10^2 \times 10^3 = 1.3 \times 10^6$;
and the cost for system 2 is $10^5 + 11 \times 10^2 \times 10^3 = 1.2 \times 10^6$.
Thus, the second one is much cheaper.

(Problem 4.2.3)

A gambler observing a game in which a single die is tossed repeatedly gets the impression that 6 comes up about 18% of the time, 5 about 14% of the time, whereas the other four numbers are equally likely to occur (i.e., with probability 0.17). Upon being asked to play, the gambler asks that he first be allowed to test his hypothesis by tossing the die n times.
(a) What test staitstic should he use if the only alternatve he considers is that the die is fair?
(b) Show that if $n = 2$ the most powerful level 0.0196 test rejects if, and only if, two $5's$ are obtained.
(c) Using the fact that if $(N_1, ..., N_k) \tilde{} M(n, \theta_1, ..., \theta_k)$, then $a_1N_1 + \cdots + a_kN_k$ has approximately a $N(n\mu, n\sigma^2)$ distribution, where $\mu = \sum_{i=1}^k a_i\theta_i$ and $\sigma^2 = \sum_{i=1}^k \theta_i (a_i - \mu)^2$, find an approximation to the critical value of the MP level α test for this problem.
Solution:

(a) To test
$H_0 : \theta_1 = 0.18, \theta_2 = 0.14, \theta_3 = \theta_4 = \theta_5 = \theta_6 = 0.17$
$vs.\ \ H_1 : \theta_1 = \theta_2 = \theta_3 = \theta_4 = \theta_5 = \theta_6 = \frac{1}{6}$,
use the LR test,

$$L(x, \theta_0, \theta_1) = \frac{P(x,\theta_1)}{P(x,\theta_2)}$$

$$= \frac{\frac{n!}{n_1!n_2!\cdots n_6!}(1/6)^n}{\frac{n!}{n_1!n_2!\cdots n_6!}0.18^{x_1}0.14^{x_2}(0.17)^{n-x_1-x_2}} = \frac{(1/6)^n}{0.18^{x_1}0.14^{x_2}(0.17)^{n-x_1-x_2}}.$$

(b) For $n = 2$,
$P(\text{two } 5's \text{ are obtained} \mid H_0) = \frac{2!}{2!}0.14^2 = 0.0196$
and $P(L(x, \theta_0, \theta_1) > k | H_0) = 0.0196$.
Then, the largest value of $L(x, \theta_0, \theta_1)$ is $(\frac{1}{0.84})^2$, which corresponds to {two 5's are obtained}.
The second largest $L(x, \theta_0, \theta_1)$ is $\left(\frac{1}{0.84}\right)\left(\frac{1}{1.02}\right)$.

As we need k between the above two values for the desired power,thus no other possible outcomes of two experiments can reject the null bypothesis with power 0.0196 other than two 5's.

(c)

$L = \frac{1}{1.08^{n_6} 0.84^{n_5} 1.02^{n-n_6-n_5}}$,

$\log L = -n_6 \log 1.08 - n_5 \log 0.84 - (n - n_6 - n_5) \log 1.02$

Then, under H_0,

$\left(n_6, n_5, \sum_{i=1}^{4} n_i\right) \tilde{} M\left(n, 0.18, 0.14, 0.68\right)$

$\Rightarrow \log L \tilde{} N\left(n\mu, n\sigma^2\right)$

Where,

$n\mu = n\sum_{i=1}^{k} a_i\theta_i, \;\; a_1 = -\log 1.08, a_2 = -\log 0.84$

$n\sigma^2 = n\sum_{i=1}^{k} \theta_i\left(a_i - \mu\right)^2, \;\; \mu = \sum_i a_i\theta_i$

So, the critical value is $\exp\left\{\sqrt{n\sigma^2} Z_{1-\alpha} + n\mu\right\}$.

7. (Optional only, recommended but not required)

(Problem 4.1.5)

Show that if H is simple and the test statistic T has a continuous distribution, then the p-value $\alpha(T)$ has a uniform, $U(0,1)$, distribution.

Hint: See Problem B.2.12.

Solution:

As H is simple,

$\alpha(k) = \underset{\Theta_0}{Sup}\left\{P_\theta\left(T > k\right)\right\} = P_{\theta_0}\left(T(x) > k\right) = 1 - F_T(k)$.

And also T is continuous, by problem B.2.12,

$F_T(k) \tilde{} U(0,1)$,

Then, *p-value* $\alpha(T)$ has a uniform, $U(0,1)$, distribution.

(Problem 4.1.6)

Suppose that $T_1, ..., T_r$ are independent test staitstics for the same simple H and that each T_j has a continuous distribution, $j = 1, ..., r$. Let $\alpha(T_j)$ denote the p-value for $T_j, j = 1, ..., r$.

(a) Show that, under $H, \hat{T} = -2\sum_{j=1}^{r} \log\alpha(T_j)$ has a χ^2_{2r} distribution.

Hint: See Problem B.3.4.

Solution:

By conclusion from 4.1.5,

$\alpha(T_j)$ has $U(0,1)$ distribution,

$\Rightarrow -\log\alpha(T_j)$ has an $Exp(1)$ distribution,

$\Rightarrow -2\log\alpha(T_j)$ has a $Gamma\left(1, \frac{1}{2}\right)$ distribution,

$\Rightarrow -2\sum_{j=1}^{r} \log\alpha(T_j)$ has a $Gamma\left(r, \frac{1}{2}\right)$ distribution,

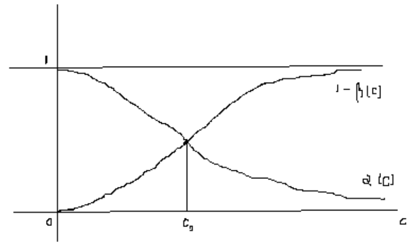
i.e. $-2\sum_{j=1}^{r} \log\alpha(T_j)$ has a χ^2_{2r} distribution.

(Problem 4.2.4)

A formulation of goodness of tests specifies that a test is best if the maximum probability of error (of either type) is as small as possible.

(a) show that if in testing $H : \theta = \theta_0$ versus $K : \theta = \theta_1$ there exists a critical value c such that

$P_{\theta_0}\left[L(X, \theta_0, \theta_1) \geq c\right] = 1 - P_{\theta_1}\left[L(X, \theta_0, \theta_1) \geq c\right]$,



then the likelihood ratio test with critical value c is best in this sense.
(b) Find the test that is best in this sense for Example 4.2.1.
Solution:
Type I error $\alpha(c) = P_{\theta_0}(L(x,\theta_0,\theta_1) \geq c)$ is decreasing in c;and Type II error $1-\beta(c) = 1 - P_{\theta_1}(L(x,\theta_0,\theta_1) \geq c)$ is increasing in c.
The curvatures corresponding to two types error will cross at c_0,i.e. c_0 is the solution of the equation
$P_{\theta_0}(L(x,\theta_0,\theta_1) \geq c) = 1 - P_{\theta_1}(L(x,\theta_0,\theta_1) \geq c)$.
And then,

$$\max(\alpha(c), 1-\beta(c)) = \begin{cases} \alpha(c) & c \leq c_0 \\ 1-\beta(c) & c \geq c_0 \end{cases}$$

$$\Rightarrow \min_c \max(\alpha(c), 1-\beta(c)) = \alpha(c_0).$$

Hence, the likelihood ratio test with critical value c_0 is best in this sense.
(b)

$$P_{\theta_0}(L(x,\theta_0,\theta_1) \geq c) = P_{\theta_0}\left(\frac{\sqrt{n}\bar{x}}{\sigma} > \frac{\sqrt{n}c}{\sigma}\right)$$

$$P_{\theta_1}(L(x,\theta_0,\theta_1) \geq c) = P_{\theta_1}\left(\frac{\sqrt{n}(\bar{x}-\nu)}{\sigma} > \frac{\sqrt{n}(c-\nu)}{\sigma}\right)$$

From

$$P_{\theta_0}(L(x,\theta_0,\theta_1) \geq c) = 1 - P_{\theta_1}(L(x,\theta_0,\theta_1) \geq c),$$

$$\Rightarrow -c = c - \nu$$

$$\Rightarrow c = \frac{\nu}{2}$$

$$\Rightarrow \delta(x) = \begin{cases} 1 & \text{if } \bar{x} > \frac{\nu}{2} \\ 0 & \text{if } \bar{x} \leq \frac{\nu}{2} \end{cases}$$

(Problem 4.2.5)
A newly discovered skull has cranial measurements (X,Y) known to be distributed either (as in population 0) according to $N(0,0,1,1,0.6)$ or (as in population 1) according to $N(1,1,1,1,0.6)$ where all parameters are known. find a statistic $T(X,Y)$ and a critical value c such that if we use the classification rule, (X,Y) belongs to population 1 if $T \geq c$, and to population o if $T \leq c$, then the maximum of the two probabilities of misclassification $P_0[T \geq c], P_1[T < c]$ is as small as possible.
Hint: Use Problem 4.2.4 and recall (Proposition B.4.2) that linear combinations of bivariate normal random variables are normallly distributed.
Solution:
As

$$H_0 : (X,Y) \tilde{} N(0,0,1,1,0.6),$$

& $H_1 : (X,Y)\,\tilde{}\,N(1,1,1,1,0.6)$;
Let $T = X + Y$, then

$$T|H_0\,\tilde{}\,N(0,3.2) \quad \& \quad T|H_1\,\tilde{}\,N(2,3.2)$$

And thus,

$$L(T,\theta_0,\theta_1) = \frac{e^{-\frac{(T-2)^2}{2(3.2)}}}{e^{-\frac{T^2}{2(3.2)}}},$$

$$\Rightarrow \log L(T,\theta_0,\theta_1) = -\frac{(T-2)^2}{6.4} + \frac{T^2}{6.4} = \frac{T-1}{1.6}$$

$$\delta(x,y) = \begin{cases} 1 & T \geq 1.6c + 1 \\ 0 & T < 1.6c + 1 \end{cases}$$

Let $k = 1.6c + 1$, then

$$\alpha(k) = P_{\theta_0}(T \geq k) = 1 - P_{\theta_1}(T \geq k) = 1 - \beta(k)$$

$$\Rightarrow 1 - \Phi\left(\frac{k}{\sqrt{3.2}}\right) = 1 - \left(1 - \Phi\left(\frac{k-2}{\sqrt{3.2}}\right)\right)$$

$$\Rightarrow 1 - \Phi\left(\frac{k}{\sqrt{3.2}}\right) = \Phi\left(\frac{k-2}{\sqrt{3.2}}\right)$$

$$\Rightarrow k = -(k-2)$$

$$\Rightarrow k = 1$$

So,

$$\delta(x,y) = \begin{cases} 1 & if\ T \geq 1 \\ 0 & if\ T < 1 \end{cases}.$$