Solutions for 100324 Inference, CB. (1)

Ex 1b) $X_1, \ldots X_n$, $P(x_i=1)=\theta$, $X_i \sim Be(\theta)$. Verify that $T=\sum x_i$ is suff. for $\theta$.

□ $$P(X_1=x_1,\ldots,X_n=x_n / T=t) = \frac{P(X_1=x_1,\ldots,X_n=x_n, T=t)}{P(T=t)} =$$

$$= \frac{\theta^t(1-\theta)^{n-t}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}} = \frac{1}{\binom{n}{t}},$$ so by def. we showed that the conditional distr. of $X_1,\ldots X_n$ given $T$ does not depend on $\theta$. Interpretation: Given the total number of ones in $n$ trials, the probability that they occur on any particular set of $t$ trials is the same.

Ex 1c) Use Thm. 6.2.B: $\frac{f(x/\theta)}{f(y/\theta)}$ is const wrt $\theta$ iff $x \equiv y$.

$$\frac{f(x/\theta)}{f(y/\theta)} = \frac{(2\pi)^{-n/2} e^{-\sum_i (x_i-\theta)^2/2}}{(2\pi)^{-n/2} e^{-\sum(y_i-\theta)^2/2}} = \exp\left(-\frac{1}{2}\left(\sum^n x_i^2 - \sum y_i^2\right) + 2\theta n(\bar{y}-\bar{x})\right)$$

This is a const. wr to $\theta$ iff $\bar{y}=\bar{x} \Rightarrow \bar{x}$ is min suff. for $\theta$

Ex. 2 a) $E[\sum a_i X_i] = \sum_i a_i E[X_i] = \mu \underbrace{\sum a_i}_{1} = \mu$ (unbiased)

b) $Var[\sum_i a_i X_i] = \sum_i a_i^2 Var[X_i] = \sum a_i^2 \sigma^2 = \sigma^2 \sum a_i^2$. So we need to minimize $\sum_i a_i^2$ subject to the constraint $\sum_i a_i = 1$. Add and subtr. the mean of $a_i$, $\frac{1}{n}$ to have

$$\sum a_i^2 = \sum_i \left[\left(a_i - \frac{1}{n}\right) + \frac{1}{n}\right]^2 = \sum_i \left(a_i - \frac{1}{n}\right)^2 + \frac{1}{n^2} \quad (\text{cross term} = 0)$$

Hence $\sum a_i^2$ is minimized by setting $a_i = \frac{1}{n}$ $\forall i = \overline{1,n}$.
Thus $\bar{X}$ has a min variance of all unbiased estimators.

Ex. 3a) $\beta(\theta) = P_\theta\left(\frac{|\bar{X}-\theta_0|}{\sigma/\sqrt{n}} > c\right) = 1 - P_\theta\left(\frac{|\bar{X}-\theta_0|}{\sigma/\sqrt{n}} \le c\right) =$

$= 1 - P_\theta\left(-\frac{c\sigma}{\sqrt{n}} \le \bar{X}-\theta_0 \le \frac{c\sigma}{\sqrt{n}}\right) =$

$= 1 - P_\theta\left(\frac{-c\sigma/\sqrt{n}+\theta_0-\theta}{\sigma/\sqrt{n}} \le \frac{\bar{X}-\theta}{\sigma/\sqrt{n}} \le \frac{c\sigma/\sqrt{n}+\theta_0-\theta}{\sigma/\sqrt{n}}\right) =$

$= 1 - P_\theta\left(-c+\frac{\theta_0-\theta}{\sigma/\sqrt{n}} \le Z \le c+\frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right) =$

$= 1 + \Phi\left(-c+\frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right) - \Phi\left(c+\frac{\theta_0-\theta}{\sigma/\sqrt{n}}\right).$

b) $0.05 = \beta(\theta_0) = 1 - \Phi(-c) - \Phi(c) \Rightarrow c = 1.96$. The power, (1 - type II error) is

$0.75 \le \beta(\theta_0+\sigma) = 1 - \Phi(-c-\sqrt{n}) - \Phi(c-\sqrt{n}) =$

$= 1 + \Phi(-1.96-\sqrt{n}) - \Phi(1.96-\sqrt{n})$.

$\Phi(-0.675) \approx 0.25 \Rightarrow 1.96-\sqrt{n} = -0.675 \Rightarrow n = 6.943 \approx 7.$

Ex. 4a) $L(\tau) = \left(\frac{1}{\tau}\right)^n e^{-\frac{1}{\tau}\sum x_i}, \ \tau > 0$

$\ell(\tau) = \log L(\tau) = -n\log\tau - \frac{1}{\tau}\sum x_i, \quad \ell'(\tau) = -\frac{n}{\tau} + \frac{1}{\tau^2}\sum^n x_i = 0 \Leftrightarrow$

$\Rightarrow \frac{1}{\tau^2}\sum x_i = \frac{n}{\tau} \Rightarrow \hat{\tau} = \bar{x}.$

$\ell''(\tau) = \frac{n}{\tau^2} - \frac{2}{\tau^3}\sum x_i = \frac{n}{\tau^2}\left(1 - \frac{2}{\tau}\bar{x}_n\right) < 0 \Rightarrow \hat{\tau}_{ML} = \bar{x}.$

b) $E[\hat{\tau}_n(x)] = E[\bar{x}_n] = \frac{1}{n}\sum E[x_i] = \tau, \ V[\hat{\tau}_n] = \frac{1}{n}V[x_i] = \frac{\tau^2}{n}.$

c) Since $E[\hat{\tau}_n] = \tau$ and $Var[\hat{\tau}_n] = \frac{\tau^2}{n}$ the CLT gives

$\frac{\bar{x}_n-\tau}{\tau/\sqrt{n}} = \frac{\hat{\tau}_n-\tau}{\tau/\sqrt{n}} \xrightarrow[n\to\infty]{d} N(0,1)$. Using the fact that

$\hat{\tau}_n \xrightarrow{P} \tau$, we can also write $\frac{\hat{\tau}_n-\tau}{\hat{\tau}_n/\sqrt{n}} \to N(0,1)$.

d) Since by the C-R inequality

$$V(\tilde{\tau}_n) \geq \frac{1}{nI(\tau)} = \frac{1}{n\left\{-E\left[\frac{\partial^2}{\partial \tau^2}\log f(x_i|\tau)\right]\right\}} = \frac{1}{nE\left[-\frac{1}{\tau^2}+\frac{2}{\tau^2}X\right]} =$$

$= \frac{1}{n/\tau^2} = \frac{\tau^2}{n}$, for any other unbiased estimator $\tilde{\tau}_n$ of $\tau$, we see that m.l.e. attains the smallest variance possible among all unbiased estimators. Thus there is no other unbiased estimator with var. $< \frac{\tau^2}{n}$.

e) Using the result c), we have

$$P\left(-z_{\alpha/2} \leq \frac{\hat{\tau}_n-\tau}{\tau/\sqrt{n}} \leq z_{\alpha/2}\right) \approx 1-\alpha \iff$$

$$\iff P\left(-\frac{z_{\alpha/2}}{\sqrt{n}} \leq \frac{\hat{\tau}_n}{\tau}-1 \leq \frac{z_{\alpha/2}}{\sqrt{n}}\right) \approx 1-\alpha \iff$$

$$\iff P\left(-\frac{z_{\alpha/2}}{\sqrt{n}}+1 \leq \frac{\hat{\tau}_n}{\tau} \leq \frac{z_{\alpha/2}}{\sqrt{n}}+1\right) = 1-\alpha \iff$$

$$\iff P\left(\frac{\hat{\tau}_n}{1+z_{\alpha/2}/\sqrt{n}} \leq \tau \leq \frac{\hat{\tau}_n}{1-z_{\alpha/2}/\sqrt{n}}\right) = 1-\alpha.$$

$$\to I_\tau = [\,;\,]$$

Ex 5. The mle of $\beta$ is $X_{(n)} = \max_i X_i$. Since $\beta$ is a scale parameter, $X_{(n)}/\beta$ is a pivot, and

$$0.05 = P_\beta\left(\frac{X_{(n)}}{\beta} \leq c\right) = P_\beta(\text{all } x_i \leq c\beta) = \left(\frac{c\beta}{\beta}\right)^{\alpha_0 n} = c^{\alpha_0 n},$$

which means that $c = (0.05)^{1/\alpha_0 n}$. Thus, $0.95 =$
$= P_\beta(X_{(n)}/\beta > c) = P(X_{(n)}/c > \beta)$, and $\left\{\beta : \beta < \frac{X_{(n)}}{0.05^{1/\alpha_0 n}}\right\}$ is a 95% upper confidence limit for $\beta$

Ex6: a, b) $P(\overline{X} > \theta_0 + z_\alpha \sigma/\sqrt{n} \mid \theta_0) = P\left(\frac{\overline{X}-\theta_0}{\sigma/\sqrt{n}} > z_\alpha \mid \theta_0\right) = P(Z > z_\alpha) = \alpha$

Since $\overline{x}$ is unrestricted maximum le, and the restricted MLE is $\theta_0$ if $\overline{x} > \theta_0$, the LRT statistic is, for $\overline{x} > \theta_0$

$$\lambda(x) = \frac{(2\pi\sigma^2)^{-n/2} e^{-\sum(x_i-\theta_0)^2/2\sigma^2}}{(2\pi\sigma^2)^{-n/2} e^{-\sum(x_i-\overline{x})^2/2\sigma^2}} =$$

$$= \frac{e^{-[n(\overline{x}-\theta_0)^2+(n-1)S^2]/2\sigma^2}}{e^{-(n-1)S^2/2\sigma^2}} = e^{-n(\overline{x}-\theta_0)^2/2\sigma^2}, \text{ and}$$

$\lambda(x)$ is 1 for $\overline{x} < \theta_0$. Thus, rejecting if $\lambda(x) < c$ is equivalent to rejecting if $\frac{\overline{x}-\theta_0}{\sigma/\sqrt{n}} > c'$.

c) Since $\overline{X}$ is the sufficient statistics and the likelihood ratio, $\lambda(x)$ is a monotone in $\overline{x}$, the test is UMP test of size $\alpha$.