STAT 723

Solutions to Assignment #3

(10.1) A sufficient condition for a sequence of estimators $\{T_n\}$ to be consistent is that $\text{Bias}^2_\theta(T_n) + \text{Var}_\theta(T_n) \to 0$ as $n \to \infty$.

There are a lot of ways to find a consistent estimator. Perhaps the easiest way is to find a function $h(X_1)$ such that $E_\theta h(X_1) = \theta$ and $\text{Var}_\theta h(X_1) < \infty$. Then $T_n = \frac{\sum_{i=1}^n h(X_i)}{n}$ is unbiased, and its variance $\frac{\text{Var}_\theta h(X_1)}{n} \to 0$ as $n \to \infty$, so $\{T_n\}$ is consistent.

(continued)

10.1 (continued)

$$E_\theta(X_1) = \int x f(x|\theta)\,dx = \int_{-1}^{1} \tfrac{1}{2} x(1+\theta x)\,dx = \tfrac{1}{2}\int_{-1}^{1} [x+\theta x^2]\,dx$$

$$= \tfrac{1}{2}\left[\frac{x^2}{2} + \theta\frac{x^3}{3}\right]\Bigg|_{-1}^{1} = \tfrac{1}{2}\left[\tfrac{1}{2}+\tfrac{\theta}{3} - \left(\tfrac{1}{2}-\tfrac{\theta}{3}\right)\right] = \tfrac{1}{2}\,\tfrac{2}{3}\theta = \tfrac{1}{3}\theta.$$

So $E_\theta(3X_1) = \theta$.

Also $\mathrm{var}(3X_1) \le 9\,\mathrm{var}(X_1) \le 9\,EX_1^2 \le 9 < \infty$

since $P_\theta[-1<X<1]=1$, for all $\theta \in (-1,1)$

So, by the argument given earlier,

$$T_n = \frac{3\sum_{i=1}^n X_i}{n} \text{ is a consistent est of } \theta.$$

(There are, of course, many other consistent estimators of θ).

10.3 (a) The likelihood function is

$$L(\theta|x_1,\dots,x_n) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n e^{-\frac{1}{2\theta}\sum(x_i-\theta)^2}$$

so $\log L(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{1}{2\theta}\sum(x_i-\theta)^2.$

Assuming that $\log L(\theta)$ is maximized at some interior point of $(0,\infty)$

(this needs to be checked), we solve $\frac{d}{d\theta}\log L(\theta) = 0$, or

$$-\frac{n}{2\theta} - \frac{1}{2\theta}\sum 2(x_i-\theta)(-1) + \frac{1}{2\theta^2}\sum(x_i-\theta)^2 = 0$$

or $-\frac{n}{2\theta} + \frac{2}{2\theta}\left(\sum x_i - n\theta\right) + \frac{1}{2\theta^2}\left[\sum x_i^2 - 2\theta\sum x_i + n\theta^2\right] = 0$

(continued)

10.3(a) (continued)

OR (multiplying by $2\theta^2 > 0$)

$$-n\theta + 2\theta\sum X_i - 2n\theta^2 + \sum X_i^2 - 2\theta\sum X_i + n\theta^2 = 0$$

OR $\quad -n\theta^2 - n\theta + \sum X_i^2 = 0$

OR $\quad \theta^2 + \theta - \frac{\sum X_i^2}{n} = 0 \quad$ OR $\quad \theta^2 + \theta - W_n = 0$, where $W_n = \frac{\sum X_i^2}{n}$.

This quadratic equation has the two real roots

$$\frac{-1 \pm \sqrt{1+4W_n}}{2}.$$

Since $\frac{1}{2}(-1-\sqrt{1+4W_n}) < 0$ is outside the parameter space $[\theta \in (0,\infty)]$, and the other solution $\frac{1}{2}[\sqrt{1+4W_n} - 1] \geq 0$, the unique local extremum attained there must necessarily be the global maximum — that is the MLE is $\hat{\theta}_n = \frac{1}{2}[\sqrt{1+4W_n} - 1]$.

(b) By the theorem on asymptotic efficiency of MLE's,

$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} N(0, v(\theta))$, where $v(\theta) = \frac{1}{I(\theta)}$.

(continued)

10.3 (b) (continued)

To calculate $I(\theta)$, note that

$$\log f(x|\theta) = \log\left(\frac{1}{\sqrt{2\pi\theta}}\right)e^{-\frac{1}{2\theta}(x-\theta)^2} = -\frac{1}{2}\log(2\pi) + \theta^{-1/2} - \frac{1}{2\theta}(x-\theta)$$

so that (after differentiation and doing some algebraic simplification)

I get that $\frac{\partial}{\partial\theta}\log f(x|\theta) = \frac{1}{2\theta^2}[x^2-\theta^2-\theta]$.

Hence $$V(\theta) = \frac{1}{E_\theta\left\{\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]^2\right\}} = \frac{4\theta^4}{E_\theta\left\{\left[X^2-(\theta^2+\theta)\right]^2\right\}}$$

$$= \frac{4\theta^2}{E_\theta\{[X^2-E_\theta X^2]^2\}} = \frac{4\theta^4}{Var_\theta X^2} = \frac{4\theta^4}{E_\theta(X^4)-(E_\theta(X^2))^2}$$

noting that
$E_\theta X^2 = Var_\theta X + (E_\theta X)^2 = \theta + \theta^2$

Some further calculation yield that the denominator $= 4\theta^3+\theta^2$,

so that $V(\theta) = \frac{4\theta^4}{4\theta^3+\theta^2} = \frac{4\theta^2}{4\theta+1}$.

So, for large n, $Var(\hat{\theta}_n) \approx \frac{4\theta^2}{n(4\theta+1)}$.

When θ is unknown, a reasonable estimate of the approximate variance of $Var(\hat{\theta}_n)$ is $\frac{4\hat{\theta}_n^2}{n(4\hat{\theta}_n+1)}$.

10.33 I already filled in the gaps by providing details of the proof of Theorem 10.3.1 in an earlier lecture.

10.34 (a) $L(\theta|\underset{\sim}{x}) = \prod_{i=1}^{n}\left[p^{x_i}(1-p)^{1-x_i}\right] = p^{\Sigma x_i}(1-p)^{n-\Sigma x_i}$

The MLE of p is easily seen to be $\hat{p} = \Sigma x_i/n$.

So $\lambda(\underset{\sim}{x}) = \dfrac{p_0^{n\hat{p}}(1-p_0)^{n(1-\hat{p})}}{\hat{p}^{n\hat{p}}(1-\hat{p})^{n(1-\hat{p})}}$

$$-2\log\lambda(\underset{\sim}{x}) = -2\left[n\hat{p}\log p_0 + n(1-\hat{p})\log(1-p_0) - n\hat{p}\log\hat{p} - n(1-\hat{p})\log(1-\hat{p})\right]$$

$$= -2n\hat{p}\log(p_0/\hat{p}) - 2n(1-\hat{p})\log\left[(1-p_0)/(1-\hat{p})\right]$$

10.35 (a) The exact distribution of $Z = \dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \dfrac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma}$ is $N(0,1)$ under $H_0: \mu = \mu_0$. Since σ^2 is known, there is no need to replace the known σ in the denominator by a consistent estimator. So the simplest "Wald" test rejects H_0 when $|Z| > z_{\alpha/2}$ (an exact test for all n).

(b) With μ known, $(X_i-\mu)^2$, $i=1,\ldots,n$, are i.i.d., each with mean σ^2 and variance $V(\sigma) = E\{[(X_1-\mu)^2 - \sigma^2]^2\}$.

So by the CLT, $\dfrac{\frac{1}{n}\sum_{i}^{n}(X_i-\mu)^2 - \sigma_0^2}{\sqrt{V(\sigma_0)/n}} \xrightarrow{\mathcal{D}} N(0,1)$ under $H_0: \sigma^2 = \sigma_0^2$;

(continued)

10.35 (b) (Continued) To compute $V(\sigma)$, write

$Y_i = X_i - \mu$, so $Y \sim N(0, \sigma^2)$

$$V(\sigma) = E\{[Y^2 - \sigma^2]^2\} = EY^4 - 2\sigma^2 EY^2 + \sigma^4$$

$$= \sigma^4 EZ^4 - 2\sigma^4 + \sigma^4 = 3\sigma^4 - 2\sigma^4 + \sigma^4 = 2\sigma^4.$$

Hence, under $H_0: \sigma^2 = \sigma_0^2$,

$$\frac{\sqrt{n}\left[\frac{1}{n}\sum_1^n (X_i - \mu)^2 - \sigma_0^2\right]}{\sqrt{2}\,\sigma_0^2} \xrightarrow{\mathcal{D}} N(0,1)$$

Since the denom. $\sqrt{2}\,\sigma_0^2$ is known under H_0, it need not be estimated. Hence an approx level $1-\alpha$ test of $H_0: \sigma = \sigma_0$ vs $H_1: \sigma \neq \sigma_0$, for large n, is to reject H_0 when the above statistic is $> z_{\alpha/2}$ or $< -z_{\alpha/2}$.

Note that this Wald test is inferior to the exact test based on the fact that $\frac{\sum (X_i - \mu)^2}{\sigma_0^2} \sim \chi_n^2$ under H_0.

10.36 (a) $$L(\beta/\underset{\sim}{x}) = \prod_{i=1}^n \left[\frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-x_i/\beta}\right]$$

$$= \frac{1}{[\Gamma(\alpha)]^n \beta^{n\alpha}} \left(\prod x_i\right)^{\alpha-1} e^{-\sum x_i/\beta}$$

So $\log L(\beta/\underset{\sim}{x}) = -n\log\Gamma(\alpha) - n\alpha\log\beta + (\alpha-1)\sum\log x_i - \frac{\sum x_i}{\beta}$

(continued)

10.36 (a) (continued)

$$\frac{d}{d\beta} \log L(\beta|\underset{\sim}{x}) = -\frac{n\alpha}{\beta} + \frac{\Sigma x_i}{\beta^2}$$

$$\frac{d^2}{d\beta^2} \log L(\beta|\underset{\sim}{x}) = \frac{n\alpha}{\beta^2} - 2\frac{\Sigma x_i}{\beta^3}$$

The solution to $\frac{d}{d\beta}\log L(\beta|\underset{\sim}{x}) = 0$ is $\beta = \frac{\Sigma x_i}{n\alpha} = \frac{\overline{X}}{\alpha}$,

and $\frac{d^2}{d\beta^2}\log L(\beta|\underset{\sim}{x})\Big|_{\beta=\frac{\overline{X}}{\alpha}} = \frac{1}{(\overline{X}/\alpha)^2}\left[n\alpha - \frac{2(\Sigma x_i)n\alpha}{\Sigma x_i}\right] < 0$,

so that $\hat{\beta} = \overline{X}/\alpha$ is the MLE of β.

(b) $E\hat{\beta} = EX_1/\alpha = \frac{\alpha\beta}{\alpha} = \beta$,

and $Var(\hat{\beta}) = \frac{1}{\alpha^2}\frac{1}{n}Var(X_1) = \frac{1}{n\alpha^2}\alpha\beta^2 = \frac{1}{n}\frac{\beta^2}{\alpha}$.

By the CLT, $\frac{\sqrt{n}(\hat{\beta}-\beta)}{\beta/\sqrt{\alpha}} \xrightarrow{\mathcal{D}} N(0,1)$ as $n\to\infty$.

Under $H_0: \beta = \beta_0$, $Z = \frac{\sqrt{n}\left(\frac{\overline{X}}{\alpha}-\beta_0\right)}{\beta_0/\sqrt{\alpha}} \xrightarrow{\mathcal{D}} N(0,1)$.

I think that a perfectly good Wald-type test of $H_0: \beta = \beta_0$ is to reject H_0 when $|Z| > z_{\alpha/2}$. However the hint in the textbook suggests replacing the β_0 in the denominator by its consistent est $\hat{\beta} = \frac{\overline{X}}{\alpha}$, yielding the test statistic

$\frac{\sqrt{n}\left(\frac{\overline{X}}{\alpha}-\beta_0\right)}{\overline{X}/\alpha^{3/2}}$ which also $\xrightarrow{\mathcal{D}} N(0,1)$ as $n\to\infty$.

10.36 (c) Writing $\sigma^2 = Var(X_1) \cdot (= \alpha\beta^2)$,

we have that $\dfrac{\sqrt{n}\left(\frac{\bar{X}}{\alpha} - \beta\right)}{\sqrt{\frac{1}{\alpha^2}\sigma^2}} \xrightarrow{\mathcal{D}} N(0,1)$.

Since all moments of a gamma r.v. are finite, we still have a limiting normal dist if the σ^2 above is replaced by its consistent est. $S^2 = \dfrac{\Sigma(X_i - \bar{X})^2}{n-1}$.

Hence, yet another version of a Wald test is to set $Z = \dfrac{\alpha\sqrt{n}\left(\frac{\bar{X}}{\alpha} - \beta\right)}{S}$ and reject H_0 if $|Z| > z_{\alpha/2}$.

10.37 (a) [μ unknown, σ^2 known]

$$S(\mu) = \frac{\partial}{\partial\mu}\left[\log\left(\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\Sigma(x_i-\mu)^2}\right)\right]$$

$$= \frac{\partial}{\partial\mu}\left[\text{const.} - \frac{1}{2\sigma^2}\Sigma(x_i-\mu)^2\right] = -\frac{1}{2\sigma^2}2\Sigma(x_i-\mu)(-1)$$

$$= \frac{1}{\sigma^2}n(\bar{x}-\mu)$$

Also $I_n(\mu) = -E_\mu\left[\frac{\partial}{\partial\mu}\frac{1}{\sigma^2}n(\bar{x}-\mu)\right] = \frac{n}{\sigma^2}$

So $Z_S = S(\mu_0)/\sqrt{I_n(\mu_0)} = \dfrac{\frac{n}{\sigma^2}(\bar{X}-\mu_0)}{\sqrt{n/\sigma^2}}$

$= \dfrac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma}$, (continued)

10.37 (b) [σ^2 unknown, μ known]

$$S(\sigma^2) = \frac{d}{d\sigma^2}\left[\log\left[\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\Sigma(X_i-\mu)^2}\right]\right]$$

$$= \frac{d}{d\sigma^2}\left[C - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\Sigma(X_i-\mu)^2\right]$$

$$= -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\Sigma(X_i-\mu)^2$$

Also

$$I_n(\sigma^2) = -E_{\sigma^2}\left[\frac{\partial}{\partial\sigma^2}\left[-\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\Sigma(X_i-\mu)^2\right]\right]$$

$$= -E_{\sigma^2}\left[\frac{n}{2(\sigma^2)^2} - \frac{2}{2(\sigma^2)^3}\Sigma(X_i-\mu)^2\right]$$

$$= -\left[\frac{n}{2(\sigma^2)^2} - \frac{2}{2(\sigma^2)^3}n\sigma^2\right] = -\frac{n}{2(\sigma^2)^2} + \frac{2n}{2(\sigma^2)^2}$$

$$= \frac{n}{2(\sigma^2)^2}$$

So $Z_S = S(\sigma_0^2)/\sqrt{I_n(\sigma_0^2)}$

$$= \left[-\frac{n}{2\sigma_0^2} + \frac{n}{2(\sigma_0^2)^2}\frac{\Sigma(X_i-\mu)^2}{n}\right] \Big/ \sqrt{\frac{n}{2(\sigma_0^2)^2}}$$

$$= \frac{n}{2(\sigma_0^2)^2}\left[\sigma_0^2 - \frac{\Sigma(X_i-\mu)^2}{n}\right] \Big/ \sqrt{\frac{n}{2(\sigma_0^2)^2}}$$

$$= \sqrt{\frac{n}{2}}\cdot\frac{1}{\sigma_0^2}\left[\sigma_0^2 - \frac{\Sigma(X_i-\mu)^2}{n}\right]$$

10.38 From problem 10.36, $\frac{\partial}{\partial \beta} L(\beta|\underset{\sim}{x}) = \frac{-n\alpha\beta + \Sigma x_i}{\beta^2}$

$$= n\alpha\left(\frac{\bar{x}}{\alpha} - \beta\right)/\beta^2$$

and $-E_\beta\left[\frac{\partial^2}{\partial\beta^2}\log(\beta|\underset{\sim}{x})\right] =$

$$= -E_\beta\left[\frac{n\alpha\beta - 2\Sigma x_i}{\beta^3}\right] = -\left[\frac{n\alpha\beta - 2n\alpha\beta}{\beta^3}\right] = \frac{n\alpha\beta}{\beta^3} = \frac{n\alpha}{\beta^2}.$$

So the score statistic for testing $H_0: \beta = \beta_0$ is

$$Z_S = S(\beta_0)/\sqrt{I_n(\beta_0)} = \frac{n\alpha\left(\frac{\bar{x}}{\alpha} - \beta_0\right)/\beta_0^2}{\sqrt{\frac{n\alpha}{\beta_0^2}}},$$

$$= \frac{\sqrt{n\alpha}\left(\frac{\bar{X}}{\alpha} - \beta_0\right)}{\beta_0}$$

10.40 $\log L(\lambda|\underset{\sim}{x}) = \log\left[e^{-n\lambda}\frac{\lambda^{\Sigma x_i}}{x_1!\cdots x_n!}\right]$

$$= -n\lambda + (\Sigma x_i)\log\lambda - \Sigma\log x_i$$

So $\frac{\partial}{\partial\lambda}\log L(\lambda|\underset{\sim}{x}) = -n + \frac{\Sigma x_i}{\lambda} = \frac{-n\lambda + n\bar{x}}{\lambda} = \frac{n(\bar{x}-\lambda)}{\lambda}$,

~~and~~ Also $\frac{\partial^2}{\partial\lambda^2}\log L(\lambda|\underset{\sim}{x}) = -\frac{\Sigma x_i}{\lambda^2}$,

so $-E_\lambda\left(\frac{\partial^2}{\partial\lambda^2}\log L(\lambda|\underset{\sim}{x})\right) = -E_\lambda\left(-\frac{\Sigma x_i}{\lambda^2}\right) = \frac{n\lambda}{\lambda^2} = \frac{n}{\lambda}$.

Thus
$$\frac{\frac{\partial}{\partial\lambda}\log L(\lambda|\underset{\sim}{x})}{\sqrt{-E_\lambda\left(\frac{\partial^2}{\partial\lambda^2}\log L(\lambda|\underset{\sim}{x})\right)}} = \frac{\frac{n(\bar{x}-\lambda)}{\lambda}}{\sqrt{n/\lambda}} = \sqrt{n/\lambda}\,(\bar{X}-\lambda)$$

$$= \frac{\bar{X}-\lambda}{\sqrt{\lambda/n}}.$$ ✓

10.47 (a) Using the fact that, for small p,

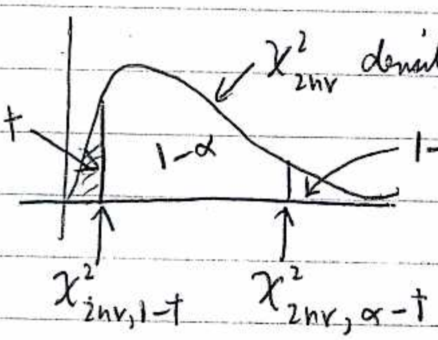
$2p \sum X_i$ is approximately distributed as χ^2_{2nr})

we have that $P_p\left[\chi^2_{2nr,1-\alpha/2} < 2p\sum X_i < \chi^2_{2nr,\alpha/2}\right] \approx 1-\alpha$

i.e. that $$P_p\left[\frac{\chi^2_{2nr,1-\alpha/2}}{2\sum X_i} < p < \frac{\chi^2_{2nr,\alpha/2}}{2\sum X_i}\right] \approx 1-\alpha$$

so that the above interval is an approx $1-\alpha$ C.I.

(b) It is clear that, for every $t \in [0,\alpha]$,



$$\left[\frac{\chi^2_{2nr,1-t}}{2\sum X_i} < p < \frac{\chi^2_{2nr,\alpha-t}}{2\sum X_i}\right]$$

is an approx $1-\alpha$ C.I.

The length of the interval is

$$\frac{1}{2\sum X_i}\left[\chi^2_{2nr,\alpha-t} - \chi^2_{2nr,1-t}\right].$$

It is clear, from the picture of the χ^2-density above, that the choice of $t \in [0,\alpha]$ which makes this length a minimum is $t=0$, with corresponding minimum length interval

$$\left[0 < p < \frac{\chi^2_{2nr,\alpha}}{2\sum X_i}\right].$$