STAT 723 Solutions to Assignment #1

(8.6) (a) The joint pdf of $X_1, \dots, X_n, Y_1, \dots, Y_m$ is

$$\frac{1}{\theta^n} e^{-\frac{1}{\theta}\Sigma x_i} \cdot \frac{1}{\mu^m} e^{-\frac{1}{\mu}\Sigma y_i} \qquad \text{all } x_i, y_i > 0.$$

The MLE's of θ and μ are $\hat{\theta} = \frac{\sum_1^n x_i}{n}$, $\hat{\mu} = \frac{\sum_1^m y_i}{m}$,

so that the denominator of the LR test statistic is

$$\frac{1}{\left(\frac{\Sigma x_i}{n}\right)^n} e^{-n} \cdot \frac{1}{\left(\frac{\Sigma y_i}{m}\right)^m} e^{-m}.$$

Under $H_0: \theta = \mu$, the MLE of $\theta (= \tau)$ is $\frac{\Sigma x_i + \Sigma y_i}{n+m}$,

so the numerator of the LR test stat is

$$\frac{1}{\left(\frac{\Sigma x_i + \Sigma y_i}{n+m}\right)^{n+m}} e^{-n-m}.$$

So the LR test rejects H_0 when $\frac{(\text{const.})(\Sigma x_i)^n (\Sigma y_i)^m}{(\Sigma x_i + \Sigma y_i)^{n+m}}$ is too small.

(b) With $T = \frac{\Sigma x_i}{\Sigma x_i + \Sigma y_i}$, the LR test statistic is clearly equivalent to $T^n(1-T)^m$. Note that since $t^n(1-t)^m$ is increasing for $t \in (0, n/(n+m))$ and decreasing for $t > n/(n+m)$, then rejecting H_0 when $T^n(1-T)^m$ is $\leq c$ is equivalent to rejecting when $T \leq c_1$ or $T \leq c_2$ for some c_1, c_2.

8.6 (c) Write $T = \dfrac{1}{1 + (\Sigma Y_i)/(\Sigma X_i)}$.

Under H_0, ΣX_i and ΣY_i are indep, with

$\dfrac{\Sigma X_i}{\theta} \sim$ gamma with $\alpha = n$, $\beta = 1$

and $\dfrac{\Sigma Y_i}{\theta} \sim$ gamma, $\alpha = m$, $\beta = 1$.

Equivalently, $\dfrac{2\Sigma X_i}{\theta} \sim \chi^2_{2n}$, $\dfrac{2\Sigma Y_i}{\theta} \sim \chi^2_{2m}$,

so that $\dfrac{\Sigma Y_i / m}{\Sigma X_i / n} \sim F_{m,n}$.

So the null distribution of $F = \dfrac{1}{1 + \frac{m}{n} T}$ is $F_{m,n}$,

or the null distribution of T is the distribution of $\left(\frac{1}{F} - 1\right)\frac{n}{m}$ where $F \sim F_{m,n}$. You can easily work out the p.d.f. of T under H_0 by making a 1-1 change of variables in the $F_{m,n}$ p.d.f.

<u>Note.</u> Writing the null density of T as $f_0(t)$, the α-level L-R tests rejects H_0 when $T < c_1$ or $T > c_2$,

where c_1 and c_2 are uniquely determined by the 2 side conditions: (i) $c_1^n (1-c_1)^m = c_2^n (1-c_2)^m$

and (ii) $\int_0^{c_1} f_0(t)\,dt + \int_{c_1}^{\infty} f_0(t) = \alpha$.

It can be shown that the two integral above are typically <u>not</u> each $= \alpha/2$.

8.8 (a) Note: in the textbook's notation $N(\theta, a\theta)$ means a normal with $\mu = \theta$ and variance $\sigma^2 = a\theta$.

The joint pdf of $X_1, \ldots X_n$ is

$$L(\theta, a) = \left(\frac{1}{\sqrt{2\pi a\theta}}\right)^n e^{-\frac{1}{2a\theta}\Sigma(X_i-\theta)^2}$$

We need to find $\hat{a}, \hat{\theta}$ which maximize

$$\log L(\theta, a) = -\frac{n}{2}\left(\log(2\pi) + \log a + \log\theta\right) - \frac{1}{2a\theta}\Sigma(X_i-\theta)^2.$$

We will solve $\frac{\partial}{\partial a}\log L = 0$, $\frac{\partial}{\partial\theta}\log L = 0$. If there is a unique soln $(\hat{\theta}, \hat{a})$, then since it (a prior) clear that the sup cannot be found ~~at the~~ as the boundary of the parameter space is approached, then the soln $(\hat{\theta}, \hat{a})$ must be a unique global maximum.

$$\frac{\partial}{\partial\theta}L = 0 \Rightarrow -\frac{n}{2\theta} - \frac{2}{2a\theta}\Sigma(X_i-\theta)(-1) + \frac{1}{2a\theta^2}\Sigma(X_i-\theta)^2 = 0$$

$$\frac{\partial}{\partial a}L = 0 \Rightarrow -\frac{n}{2a} + \frac{1}{2a^2\theta}\Sigma(X_i-\theta)^2 = 0.$$

The solution to the 2nd equation is $a = \frac{\Sigma(X_i-\theta)^2}{n\theta}$

Substituting this into the first equation yield $\theta = \bar{x}$.

Hence the MLE of (θ, a) is $\left(\bar{x}, \frac{\Sigma(X_i-\hat{\theta})^2}{n\cancel{\bar{x}}}\right) = \left(\bar{x}, \frac{S^2}{\bar{x}}\right)$.

So the denom. of the LR statistic is

$$\frac{1}{\sqrt{2\pi(S^2/\bar{x})\bar{x}}} e^{-\frac{1}{2(S^2/\bar{x})\bar{x}} nS^2} = \frac{1}{\sqrt{2\pi S^2}} e^{-n/2}$$

(continued)

8.8 (a) (continued)

Under $H_0: a=1$, the log likelihood function is

$$\log L(\theta) = -\frac{n}{2}[\log(2\pi) + \log\theta] - \frac{1}{2\theta}\sum(x_i-\theta)^2$$

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

or $\sum(x_i-\theta)^2 + 2\theta(\sum x_i - n\theta) - n\theta = 0$

or

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

or

$$n\theta^2 + n\theta - \frac{\sum x_i^2}{2} = 0, \text{ so } \theta = \frac{-1 \pm \sqrt{1+4(\sum x_i^2)/n}}{2}.$$

The restriction $\theta > 0$ (so Var >0) forces the MLE of θ under H_0 to be the positive root:

$$\hat\theta = -1 + \sqrt{1+4(\sum x_i)^2/n}.$$

So the LR test rejects H_0 when λ is too small, where

$$\lambda = \frac{\left(\frac{1}{\sqrt{2\pi\hat\theta}}\right)^n e^{-\frac{1}{2\hat\theta}\sum(x_i-\hat\theta)^2}}{\left(\frac{1}{\sqrt{2\pi s^2}}\right)^n e^{-n/2}}$$

where $\hat\theta$ is given above. This test statistic can be simplified a bit.

(b) This is similar to (a) (and probably just as tedious) so I am omitting the solution.

(5)

8.15 By the N-P lemma, ~~...~~ MP test rejects H_0 when

$$\left[\left(\frac{1}{\sqrt{2\pi\sigma_1^2}}\right)^n e^{-\frac{1}{2\sigma_1^2}\sum X_i^2}\right] \Big/ \left(\frac{1}{\sqrt{2\pi\sigma_0^2}}\right)^n e^{-\frac{1}{2\sigma_0^2}\sum X_i^2} \text{ too large,}$$

Equivalently, reject H_0 when $e^{\left(-\frac{1}{2\sigma_1^2}+\frac{1}{2\sigma_0^2}\right)\sum X_i^2}$ is too large.

Since $\sigma_0 < \sigma_1 \Rightarrow -\frac{1}{2\sigma_1^2}+\frac{1}{2\sigma_0^2} > 0$,

the MP test rejects when $\sum X_i$ is too large.

Under H_0, $\frac{\sum X_i^2}{\sigma_0^2} \sim \chi^2_n$, so $P_0\left[\frac{\sum X_i^2}{\sigma_0^2} > \chi^2_{n,\alpha}\right] = \alpha$

So the MP level α test rejects H_0 when $\sum X_i^2 > \sigma_0^2 \chi^2_{n,\alpha}$.

8.18 The power against θ is

$$P_\theta\left[\frac{\bar{X}-\theta_0}{\sigma/\sqrt{n}} > c \text{ or } < -c\right] = P_\theta\left[\frac{\bar{X}-\theta}{\sigma/\sqrt{n}} + \frac{(\theta-\theta_0)}{\sigma/\sqrt{n}} > c \text{ or } < -c\right]$$

$$= P\left[\underbrace{Z}_{N(0,1)} > c - \frac{(\theta-\theta_0)}{\sigma/\sqrt{n}}\right] + P\left[Z < -c - \frac{(\theta-\theta_0)}{\sigma/\sqrt{n}}\right]$$

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

(b) $\alpha = .05 \Rightarrow c = z_{.025} = 1.960$. at $\theta = \theta_0 + \sigma$, $\frac{\theta-\theta_0}{\sigma} = 1$.
We need to find the smallest value of n such that

$$\Phi(1.96 - \sqrt{n}) - \Phi(-1.96 - \sqrt{n}) \leq 0.25$$

I wrote a little program to compute this, and found $n = 7$ yields Type II error prob of 0.2464.

8.20 Worked out in class.

8.22 (a) By N-P lemma, MP test rejects H_0 when

$$\frac{\left(\frac{1}{4}\right)^{\Sigma x_i}\left(\frac{3}{4}\right)^{10-\Sigma x_i}}{\left(\frac{1}{2}\right)^{\Sigma x_i}\left(\frac{1}{2}\right)^{10-\Sigma x_i}} = \frac{\left(\frac{3}{4}\right)^{10}\left(\frac{1}{3}\right)^{\Sigma x_i}}{\left(\frac{1}{2}\right)^{10}}$$ is too large.

Equivalently, reject H_0 when $\sum_{i=1}^{10} x_i$ is too small.

By trial and error $P_{1/2}\left[\Sigma x_i \le 2\right] = \frac{56}{1024} = .0547$.

So MP size 0.0547 test is: Rej. H_0 if $\sum_{i=1}^{n} X \le 2$.

The power of this test is $P_{1/4}\left[X \le 2\right] = \left(\frac{3}{4}\right)^{10} + 10\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^9 + 45\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^8$
$=$ (do the arithmetic).

(b) Size $\alpha = P_{1/2}\left[\sum_{i=1}^{10} x_i \ge 6\right] = \sum_{x=6}^{10} \binom{10}{x}\left(\frac{1}{2}\right)^{10}$

$= \frac{1}{1024}\left[1+10+45+120+210\right] = \frac{386}{1024}$

(c) Exact non-randomized α-level tests exist when $\alpha = P_{1/2}\left[\sum_{i=1}^{10} x_i \le c\right]$

for some c. That is the possible α's are

$0, \frac{1}{1024}, \frac{11}{1024}, \frac{56}{1024}, \frac{176}{1024}, \ldots\ldots, \frac{1023}{1024}$ and 1.

(8.25) (a) $X \sim N(\theta, \sigma^2)$. Fix $\theta_0 < \theta_1$. Then $f(x|\theta_1)/f(x|\theta_0) =$

$$= e^{-\frac{1}{2\sigma^2}\left[(x-\theta_1)^2-(x-\theta_0)^2\right]} = \exp\left\{-\frac{1}{2\sigma^2}\left[x^2-2\theta_1 x+\theta_1^2-x^2+2\theta_0 x+\theta_0^2\right]\right\}$$

$$= e^{-\frac{1}{2\sigma^2}(\theta_1^2-\theta_0)^2} \; e^{\frac{1}{\sigma^2}(\theta_1-\theta_0)x}$$, which is monotone incr. in x.

(b) $\theta_0 < \theta_1$, then $f(x|\theta_1)/f(x|\theta_0) = e^{-\theta_1}\frac{\theta_1^x}{x!}\Big/\left[e^{-\theta_0}\theta_0^x/x!\right]$

$$= e^{-\theta_1+\theta_0}\left(\frac{\theta_1}{\theta_0}\right)^x$$, which is increasing in x.

(c) $\theta_0 < \theta_1$, then $f(x|\theta_1)/f(x|\theta_0) = \binom{n}{x}\theta_1^x(1-\theta_1)^{n-x}\Big/\left[\binom{n}{x}\theta_0^x(1-\theta_0)^{n-x}\right]$

$$= \left(\frac{1-\theta_1}{1-\theta_0}\right)^n\left[\frac{\theta_1}{1-\theta_1}\Big/\frac{\theta_0}{1-\theta_0}\right]^x$$, which is increasing in x.

(8.28) (a) Fix $\theta_0 < \theta_1$. Then $\frac{f(x|\theta_1)}{f(x|\theta_0)}$

$$= \frac{e^{(x-\theta_1)}}{\left(1+e^{(x-\theta_1)}\right)^2}\,\frac{\left(1+e^{(x-\theta_0)}\right)^2}{e^{x-\theta_0}} = e^{\theta_0-\theta_1}\left[\frac{1+e^{(x-\theta_0)}}{1+e^{(x-\theta_1)}}\right]^2.$$

Since $\frac{1+e^{(x-\theta_0)}}{1+e^{(x-\theta_1)}}$ is positive for all x, it suffices to show that $\frac{d}{dx}\left[\frac{1+e^{(x-\theta_0)}}{1+e^{(x-\theta_1)}}\right] \geq 0$ for all x.

The above derivative has the same sign as

$$\left[1+e^{(x-\theta_1)}\right]e^{(x-\theta_0)} - \left(1+e^{(x-\theta_0)}\right)e^{(x-\theta_1)}$$

$$= e^{(x-\theta_0)} - e^{(x-\theta_1)} = e^x e^{(-\theta_0+\theta_1)} > 0 \text{ for all } x.$$

Thus the family has MLR in x. (continued)

8-28 (continued)

(b) Since the family has MLR, the MP size α test must have rejection region of the form $X > c$.

$$P_0[X>c] = \int_c^\infty \frac{e^x}{(1+e^x)^2}\,dx \overset{v=1+e^{-x}}{=} \int_{1+e^c}^\infty \frac{dv}{v^2}$$

$$= -\frac{1}{v}\Big]_{1+e^c}^\infty = \frac{1}{1+e^c} = \alpha.$$

So $c = \log\left(\frac{1-\alpha}{\alpha}\right)$, and MP level α test rejects H_0

when $X > \log\left(\frac{1-\alpha}{\alpha}\right)$.

When $\alpha = .2$, reject when $X > \log 4 = 1.3863$.

$$P_{\theta=1}[\text{Reject } H_0] = \int_{\log 4}^\infty \frac{e^{(x-1)}}{(1+e^{(x-1)})^2}\,dx = \int_{y=\log 4-1}^\infty \frac{e^y}{(1+e^y)^2}\,dy$$

$$= \frac{1}{1+e^{(\log 4-1)}} = \frac{1}{1+4/e} = 0.4046$$

So the Type II error probability is $1 - 0.4046 = 0.5954$.

(c) One of the theorems was that when a family has MLR in X, then the UMP size α test of $H_0: \theta \le \theta_0$ vs $H_1: \theta < \theta_0$ has rejection region of form $X > c$, with c determined by $P_{\theta_0}[X>c] = \alpha$.

8-29 Solved in class

8.30 Solved in class.

8.31 (a) For $\lambda_0 < \lambda_1$,

$$\prod_{i=1}^{n} f(x_i|\lambda_1) \Big/ \prod_{i=1}^{n} f(x_i|\lambda_0) = e^{-n(\lambda_1-\lambda_0)} \left(\frac{\lambda_1}{\lambda_0}\right)^{\Sigma x_i}$$

is monotone increasing in Σx_i. Since this family therefore has MLR, it follows that the test which rejects H_0 when $\sum_{i=1}^{n} X_i \geq c$ is UMP level α for testing $H_0 : \lambda \leq \lambda_0$ vs. $H_1 : \lambda > \lambda_0$, where $\alpha = P_{\lambda_0}[\Sigma X_i \geq c]$.

(b) ~~[illegible]~~ $\sum_{i=1}^{n} X_i \sim$ Poisson $(n\lambda)$. When n is large, by the CLT, $\dfrac{\Sigma X_i - n\lambda}{\sqrt{n\lambda}}$ is approx $\sim N(0,1)$. So $P_{\lambda=1}[\Sigma X_i \geq c] =$

$$P_{\lambda=1}\left[\frac{\Sigma X_i - n}{\sqrt{n}} \geq \frac{c-n}{\sqrt{n}}\right] \approx P\left[Z \geq \frac{c-n}{\sqrt{n}}\right] = .05$$

This will be achieved by taking $\dfrac{c-n}{\sqrt{n}} \approx 1.645$.

Similarly to achieve

$$P_{\lambda=2}[\Sigma X_i \geq c] = P\left[\frac{\Sigma X_i - 2n}{\sqrt{2n}} \geq \frac{c-2n}{\sqrt{2n}}\right] \approx 0.9,$$

set $\dfrac{c-2n}{\sqrt{2n}} \approx -1.282$.

We now solve these two equations for n.

(continued)

8.31 (b) continued

$$c - n \approx 1.645\sqrt{n}$$

and ~~$c-2n \approx 1.645\sqrt{n}$~~

$$c - 2n \approx -1.282\sqrt{2}\,\sqrt{n}.$$

So $n \approx (1.645 + 1.282\sqrt{2})\sqrt{n}$

or $n \approx (1.645 + 1.282\sqrt{2})^2 = 11.95$

So take $n = 12$. (Note: this n is too small for the normal approx. to be very accurate. Calculations with exact Poisson probabilities may lead to a somewhat different value of the required n.)

8.33 (a) Shown in class: $k = 1 - \alpha^{1/2}$.

(b) For $0 < \theta < k$, $P_\theta[Y_n \geq 1 \text{ or } Y_1 \leq k] =$

$1 - P_\theta[Y_n < 1 \text{ and } Y_1 \leq k] =$

$$= 1 - \int_{y_1=\theta}^{k}\left[\int_{y_n=y_1}^{1} n(n-1)(y_n - y_1)^{n-2}\,dy_n\right]dy_1$$

$$= 1 - \int_{y_1=\theta}^{k}\left[\int_0^{1-y_1} n(n-1)z^{n-2}\,dz\right]dy_1$$

$$= 1 - \int_{y_1=\theta}^{k}\left[\int_0^{1-y_1} n\,dz^{n-1}\right]dy_1 = 1 - \int_{y_1=\theta}^{k} n(1-y_1)^{n-1}\,dy_1$$

$$= 1 - \int_{y_1=\theta}^{k}\left[-d(1-y_1)^n\right] = 1 - (1-\theta_1)^n - (1-k)^n$$

(8.33) (b) (continued)

So the power function is

$$\beta(\theta) = 1 - \left[(1-\theta)^n - (1-k)^n\right] \quad \left(= 1 - \left[(1-\theta)^n - \alpha\right]\right), \quad \text{for } \theta \in [0,k]$$

$$= 1 \quad , \text{ for } \theta \geq k.$$

(c) The test is VMP against $\theta \geq k$, since its power is 1 there.

So fix θ, $0 < \theta < k$. It remains to show that the test is a MP level α test of $H_0: \theta = 0$ against the simple alternative θ. We will apply the Neyman-Pearson Lemma. Recall that a MP (non-randomized) test must have the form:

$$\phi(\underset{\sim}{x}) = 1 \quad \text{if } f(\underset{\sim}{x}|\theta)/f(\underset{\sim}{x}|0) > k$$

$$= \text{"anything"}, \text{ if } f(\underset{\sim}{x}|\theta)/f(\underset{\sim}{x}|0) = k$$

$$= 0 \quad , \quad \text{if } f(\underset{\sim}{x}|\theta)/f(\underset{\sim}{x}|0) < k,$$

where "anything" means any function with range $\{\{0\},\{1\}\}$, as long as the side condition $E_0\, \phi(\underset{\sim}{X}) = \alpha$ is satisfied.

Write $r(\underset{\sim}{x}) = f(\underset{\sim}{x}|\theta)/f(\underset{\sim}{x}|0)$. Here is a complete listing of the possible values of $r(\underset{\sim}{x})$, excluding the cases $Y_1 < \theta$ and $Y_n > \theta + 1$ which are irrelevant since their pdf's are 0 under both $H_0: 0$ and $H_1: \theta$.

(continued)

8.33 (b) (continued)

event	$r(\underset{\sim}{x}) = f(\underset{\sim}{x} \mid \theta) / f(\underset{\sim}{x} \mid 0)$
$0 < y_1 < y_n < \theta$	0
$0 < y_1 < \theta < y_n < 1$	0
$0 < y_1 < \theta,\ 1 < y_n < \theta + 1$	$0/0$ (undefined, but irrelevant)
$0 < y_1 < y_n < 1$	1
$\theta < y_1 < 1 < y_n < \theta + 1$	$1/0 = \infty$
$1 < y_1 < y_n < \theta + 1$	$1/0 = \infty$

From the above, it is clear that, when $0 < \alpha < 1$, the value of k in the N-P form of ~~the~~ a MP test must be $k = 1$. Hence any MP size α test of 0 vs. θ must have the following form (where we omit all references to ranges of values of $\underset{\sim}{x}$ for which the ratio is $0/0$):

$$
\begin{aligned}
\phi(\underset{\sim}{x}) &= 1 \quad &&, \text{ if } y_1 > \theta \text{ and } y_n > 1 \\
&= \text{"anything"}, &&\text{ if } \theta < y_1 < y_n < 1 \\
&= 0 \quad &&, \text{ if } y_1 < \theta \text{ and } y_n < 1,
\end{aligned}
$$

where "anything" means any 0-1 valued function, subject to the restriction $E_0 \phi(\underset{\sim}{X}) = \alpha$.

(8.33) (c) Finally note the test under consideration **does** have the above form for each θ in $[0, k)$. It follows that the test is UMP size α for $H_0: \theta = 0$ vs. $H_1: \theta > 0$.

(d) With $n = 1$ and $k = .9$, the UMP test has level $\alpha = .10$ and power 1 against each $\theta > 1$.

(8.38) (a) The distribution of $(\bar{X} - \theta_0)/\sqrt{S^2/n}$ under $H_0: \theta = \theta_0$ is t_{n-1}. Hence

$$P_{\theta_0}\left[|\bar{X} - \theta_0| > t_{n-1, \alpha/2}\sqrt{S^2/n}\right] = P_{\theta_0}\left[\frac{\bar{X} - \theta_0}{S/\sqrt{n}} > t_{n-1, \alpha/2} \text{ or } < -t_{n-1, \alpha/2}\right]$$

$$= \alpha/2 + \alpha/2 = \alpha.$$

(b) To show that this test ("one-sample t-test") is a LRT is straighforward but tedious. For the details of the solution, see Hogg and Craig "Introduction to Mathematical Statistics", 5th ed., Example 1 on pages 413–415.

(8.41) Derivation of this LRT ("the two-sample t-test") is also straightforward but tedious. See Example 2, pages 416–419 of Hogg and Craig (above).