Stat Inf HW1 solutions 2/19/08

5.6 a. For $Z = X - Y$, set $W = X$. Then $Y = W - Z$, $X = W$, and $|J| = \begin{vmatrix} 0 & 1 \\ -1 & 1 \end{vmatrix} = 1$. Then $f_{Z,W}(z,w) = f_X(w) f_Y(w - z) \cdot 1$, thus $f_Z(z) = \int_{-\infty}^{\infty} f_X(w) f_Y(w-z) dw$.

b. For $Z = XY$, set $W = X$. Then $Y = Z/W$ and $|J| = \begin{vmatrix} 0 & 1 \\ 1/w & -z/w^2 \end{vmatrix} = -1/w$. Then $f_{Z,W}(z,w) = f_X(w) f_Y(z/w) \cdot |-1/w|$, thus $f_Z(z) = \int_{-\infty}^{\infty} |-1/w|\, f_X(w) f_Y(z/w) dw$.

c. For $Z = X/Y$, set $W = X$. Then Y=W/Z and $|J| = \begin{vmatrix} 0 & 1 \\ -w/z^2 & 1/z \end{vmatrix} = w/z^2$. Then $f_{Z,W}(z,w) = f_X(w) f_Y(w/z) \cdot |w/z^2|$, thus $f_Z(z) = \int_{-\infty}^{\infty} |w/z^2| f_X(w) f_Y(w/z) dw$.

10

c. Use the fact that $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$ and $\mathrm{Var}\chi^2_{n-1} = 2(n-1)$ to get

$$\mathrm{Var}\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1)$$

which implies $(\frac{(n-1)^2}{\sigma^4})\mathrm{Var}S^2 = 2(n-1)$ and hence

$$\mathrm{Var}S^2 = \frac{2(n-1)}{(n-1)^2/\sigma^4} = \frac{2\sigma^4}{n-1}.$$

Remark: Another approach to b), not using the χ^2 distribution, is to use linear model theory. For any matrix A $\mathrm{Var}(X'AX) = 2\mu_2^2 \mathrm{tr}A^2 + 4\mu_2\theta' A\theta$, where μ_2 is σ^2, $\theta = \mathrm{E}X = \mu 1$. Write $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}) = \frac{1}{n-1} X'(I - \bar{J}_n)X$ .Where

$$I - \bar{J}_n = \begin{pmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & & \vdots \\ \vdots & & \ddots & \vdots \\ -\frac{1}{n} & \cdots & \cdots & 1-\frac{1}{n} \end{pmatrix}.$$

Notice that $\mathrm{tr}A^2 = \mathrm{tr}A = n-1$, $A\theta = 0$. So

$$\mathrm{Var}S^2 = \frac{1}{(n-1)^2}\mathrm{Var}(X'AX) = \frac{1}{(n-1)^2}\left(2\sigma^4(n-1) + 0\right) = \frac{2\sigma^4}{n-1}.$$

5.15 a.

$$\bar{X}_{n+1} = \frac{\sum_{i=1}^{n+1} X_i}{n+1} = \frac{X_{n+1} + \sum_{i=1}^{n} X_i}{n+1} = \frac{X_{n+1} + n\bar{X}_n}{n+1}.$$

b.

$$\begin{aligned}
nS_{n+1}^2 &= \frac{n}{(n+1)-1}\sum_{i=1}^{n+1}\left(X_i - \bar{X}_{n+1}\right)^2 \\
&= \sum_{i=1}^{n+1}\left(X_i - \frac{X_{n+1}+n\bar{X}_n}{n+1}\right)^2 \qquad \text{(use (a))} \\
&= \sum_{i=1}^{n+1}\left(X_i - \frac{X_{n+1}}{n+1} - \frac{n\bar{X}_n}{n+1}\right)^2 \\
&= \sum_{i=1}^{n+1}\left[\left(X_i - \bar{X}_n\right) - \left(\frac{X_{n+1}}{n+1} - \frac{\bar{X}_n}{n+1}\right)\right]^2 \qquad (\pm\bar{X}_n) \\
&= \sum_{i=1}^{n+1}\left[\left(X_i - \bar{X}_n\right)^2 - 2\left(X_i - \bar{X}_n\right)\left(\frac{X_{n+1}-\bar{X}_n}{n+1}\right) + \frac{1}{(n+1)^2}\left(X_{n+1}-\bar{X}_n\right)^2\right] \\
&= \sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 + \left(X_{n+1}-\bar{X}_n\right)^2 - 2\frac{\left(X_{n+1}-\bar{X}_n\right)^2}{n+1} + \frac{n+1}{(n+1)^2}\left(X_{n+1}-\bar{X}_n\right)^2 \\
&\qquad \left(\text{since} \sum_1^n (X_i - \bar{X}_n) = 0\right) \\
&= (n-1)S_n^2 + \frac{n}{n+1}\left(X_{n+1}-\bar{X}_n\right)^2.
\end{aligned}$$

5.16 a. $\sum_{i=1}^{3}\left(\frac{X_i - i}{i}\right)^2 \sim \chi_3^2$

b. $\left(\frac{X_i-1}{i}\right) \Big/ \sqrt{\sum_{i=2}^{3}\left(\frac{X_i-i}{i}\right)^2 \Big/ 2} \sim t_2$

c. Square the random variable in part b).

(in 5.16b above, term should be (X1-1)/1, not (Xi-1)/i.

17

c. Write $X = \frac{U/p}{V/p}$ then $\frac{1}{X} = \frac{V/q}{U/p} \sim F_{q,p}$, since $U \sim \chi_p^2$, $V \sim \chi_q^2$ and U and V are independent.

d. Let $Y = \frac{(p/q)X}{1+(p/q)X} = \frac{pX}{q+pX}$, so $X = \frac{qY}{p(1-Y)}$ and $\left|\frac{dx}{dy}\right| = \frac{q}{p}(1-y)^{-2}$. Thus, Y has pdf

$$\begin{aligned}
f_Y(y) &= \frac{\Gamma\left(\frac{q+p}{2}\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)}\left(\frac{p}{q}\right)^{\frac{p}{2}} \frac{\left(\frac{qy}{p(1-y)}\right)^{\frac{p-2}{2}}}{\left(1+\frac{p}{q}\frac{qy}{p(1-y)}\right)^{\frac{p+q}{2}}} \frac{q}{p(1-y)^2} \\
&= \left[B\left(\frac{p}{2},\frac{q}{2}\right)\right]^{-1} y^{\frac{p}{2}-1}(1-y)^{\frac{q}{2}-1} \sim \text{beta}\left(\frac{p}{2},\frac{q}{2}\right).
\end{aligned}$$

5.21 Let m denote the median. Then, for general n we have

$$\begin{aligned}
P\left(\max(X_1,\ldots,X_n) > m\right) &= 1 - P\left(X_i \le m \text{ for } i = 1,2,\ldots,n\right) \\
&= 1 - \left[P\left(X_1 \le m\right)\right]^n = 1 - \left(\frac{1}{2}\right)^n.
\end{aligned}$$

5.24 Use $f_X(x) = 1/\theta$, $F_X(x) = x/\theta$, $0 < x < \theta$. Let $Y = X_{(n)}$, $Z = X_{(1)}$. Then, from Theorem 5.4.6,

$$f_{Z,Y}(z,y) = \frac{n!}{0!(n-2)!0!}\frac{1}{\theta}\frac{1}{\theta}\left(\frac{z}{\theta}\right)^0\left(\frac{y-z}{\theta}\right)^{n-2}\left(1-\frac{y}{\theta}\right)^0 = \frac{n(n-1)}{\theta^n}(y-z)^{n-2},\ 0<z<y<\theta.$$

Now let $W = Z/Y$, $Q = Y$. Then $Y = Q$, $Z = WQ$, and $|J| = q$. Therefore

$$f_{W,Q}(w,q) = \frac{n(n-1)}{\theta^n}(q-wq)^{n-2}q = \frac{n(n-1)}{\theta^n}(1-w)^{n-2}q^{n-1},\ 0<w<1, 0<q<\theta.$$

The joint pdf factors into functions of w and q, and, hence, W and Q are independent.

5.27 a. $f_{X_{(i)}|X_{(j)}}(u|v) = f_{X_{(i)},X_{(j)}}(u,v)/f_{X_{(j)}}(v)$. Consider two cases, depending on which of i or j is greater. Using the formulas from Theorems 5.4.4 and 5.4.6, and after cancellation, we obtain the following.

(i) If $i < j$,

$$\begin{aligned} f_{X_{(i)}|X_{(j)}}(u|v) &= \frac{(j-1)!}{(i-1)!(j-1-i)!}f_X(u)F_X^{i-1}(u)[F_X(v)-F_X(u)]^{j-i-1}F_X^{1-j}(v) \\ &= \frac{(j-1)!}{(i-1)!(j-1-i)!}\frac{f_X(u)}{F_X(v)}\left[\frac{F_X(u)}{F_X(v)}\right]^{i-1}\left[1-\frac{F_X(u)}{F_X(v)}\right]^{j-i-1}, \quad u<v. \end{aligned}$$

Note this interpretation. This is the pdf of the ith order statistic from a sample of size $j-1$, from a population with pdf given by the truncated distribution, $f(u) = f_X(u)/F_X(v)$, $u < v$.

(ii) If $j < i$ and $u > v$,

$$\begin{aligned} & f_{X_{(i)}|X_{(j)}}(u|v) \\ &= \frac{(n-j)!}{(n-1)!(i-1-j)!}f_X(u)\left[1-F_X(u)\right]^{n-i}\left[F_X(u)-F_X(v)\right]^{i-1-j}\left[1-F_X(v)\right]^{j-n} \\ &= \frac{(n-j)!}{(i-j-1)!(n-i)!}\frac{f_X(u)}{1-F_X(v)}\left[\frac{F_X(u)-F_X(v)}{1-F_X(v)}\right]^{i-j-1}\left[1-\frac{F_X(u)-F_X(v)}{1-F_X(v)}\right]^{n-i}. \end{aligned}$$

This is the pdf of the $(i-j)$th order statistic from a sample of size $n-j$, from a population with pdf given by the truncated distribution, $f(u) = f_X(u)/(1-F_X(v))$, $u > v$.

b. From Example 5.4.7,

$$f_{V|R}(v|r) = \frac{n(n-1)r^{n-2}/a^n}{n(n-1)r^{n-2}(a-r)/a^n} = \frac{1}{a-r}, \quad r/2 < v < a-r/2.$$

5.30 From the CLT we have, approximately, $\bar{X}_1 \sim \mathrm{n}(\mu, \sigma^2/n)$, $\bar{X}_2 \sim \mathrm{n}(\mu, \sigma^2/n)$. Since $\bar{X}_1$ and $\bar{X}_2$ are independent, $\bar{X}_1 - \bar{X}_2 \sim \mathrm{n}(0, 2\sigma^2/n)$. Thus, we want

$$\begin{aligned} .99 &\approx P\left(\left|\bar{X}_1-\bar{X}_2\right| < \sigma/5\right) \\ &= P\left(\frac{-\sigma/5}{\sigma/\sqrt{n/2}} < \frac{\bar{X}_1-\bar{X}_2}{\sigma/\sqrt{n/2}} < \frac{\sigma/5}{\sigma/\sqrt{n/2}}\right) \\ &\approx P\left(-\frac{1}{5}\sqrt{\frac{n}{2}} < Z < \frac{1}{5}\sqrt{\frac{n}{2}}\right), \end{aligned}$$

where $Z \sim \mathrm{n}(0,1)$. Thus we need $P(Z \ge \sqrt{n}/5(\sqrt{2})) \approx .005$. From Table 1, $\sqrt{n}/5\sqrt{2} = 2.576$, which implies $n = 50(2.576)^2 \approx 332$.

5.34 Using $\mathrm{E}\bar{X}_n = \mu$ and $\mathrm{Var}\bar{X}_n = \sigma^2/n$, we obtain

$$\mathrm{E}\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma} = \frac{\sqrt{n}}{\sigma}\mathrm{E}(\bar{X}_n-\mu) = \frac{\sqrt{n}}{\sigma}(\mu-\mu) = 0.$$

$$\mathrm{Var}\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma} = \frac{n}{\sigma^2}\mathrm{Var}(\bar{X}_n-\mu) = \frac{n}{\sigma^2}\mathrm{Var}\bar{X} = \frac{n}{\sigma^2}\frac{\sigma^2}{n} = 1.$$