Chapter 5

Properties of a Random Sample

5.1 Let $X = \#$ color blind people in a sample of size n. Then $X \sim \text{binomial}(n,p)$, where $p = .01$. The probability that a sample contains a color blind person is $P(X > 0) = 1 - P(X = 0)$, where $P(X=0) = \binom{n}{0}(.01)^0(.99)^n = .99^n$. Thus,

$$P(X > 0) = 1 - .99^n > .95 \Leftrightarrow n > \log(.05)/\log(.99) \approx 299.$$

5.3 Note that $Y_i \sim$ Bernoulli with $p_i = P(X_i \geq \mu) = 1 - F(\mu)$ for each i. Since the Y_i's are iid Bernoulli, $\sum_{i=1}^n Y_i \sim \text{binomial}(n, p = 1 - F(\mu))$.

5.5 Let $Y = X_1 + \cdots + X_n$. Then $\bar{X} = (1/n)Y$, a scale transformation. Therefore the pdf of $\bar{X}$ is $f_{\bar{X}}(x) = \frac{1}{1/n} f_Y\left(\frac{x}{1/n}\right) = n f_Y(nx)$.

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 $f_Z(z) = \int_{-\infty}^{\infty} |w/z^2| f_X(w) f_Y(w/z)dw$.

5.7 It is, perhaps, easiest to recover the constants by doing the integrations. We have

$$\int_{-\infty}^{\infty} \frac{B}{1+\left(\frac{\omega}{\sigma}\right)^2} d\omega = \sigma\pi B, \qquad \int_{-\infty}^{\infty} \frac{D}{1+\left(\frac{\omega-z}{\tau}\right)^2} d\omega = \tau\pi D$$

and

$$\begin{aligned} &\int_{-\infty}^{\infty} \left[\frac{A\omega}{1+\left(\frac{\omega}{\sigma}\right)^2} - \frac{C\omega}{1+\left(\frac{\omega-z}{\tau}\right)^2}\right] d\omega \\ &= \int_{-\infty}^{\infty} \left[\frac{A\omega}{1+\left(\frac{\omega}{\sigma}\right)^2} - \frac{C(\omega-z)}{1+\left(\frac{\omega-z}{\tau}\right)^2}\right] d\omega - Cz \int_{-\infty}^{\infty} \frac{1}{1+\left(\frac{\omega-z}{\tau}\right)^2} d\omega \\ &= A\frac{\sigma^2}{2} \log\left[1+\left(\frac{\omega}{\sigma}\right)^2\right] - \frac{C\tau^2}{2} \log\left[1+\left(\frac{\omega-z}{\tau}\right)^2\right]\Bigg|_{-\infty}^{\infty} - \tau\pi C z. \end{aligned}$$

The integral is finite and equal to zero if $A = M\frac{2}{\sigma^2}$, $C = M\frac{2}{\tau^2}$ for some constant M. Hence

$$f_Z(z) = \frac{1}{\pi^2\sigma\tau}\left[\sigma\pi B - \tau\pi D - \frac{2\pi M z}{\tau}\right] = \frac{1}{\pi(\sigma+\tau)} \, \frac{1}{1+\left(z/(\sigma+\tau)\right)^2},$$

if $B = \frac{\tau}{\sigma+\tau}$, $D = \frac{\sigma}{\sigma+\tau)}$, $M = \frac{-\sigma\tau^2}{2z(\sigma+\tau)} \frac{1}{1+\left(\frac{z}{\sigma+\tau}\right)^2}$.

$$= \frac{1}{\sqrt{2\pi}} \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2}} \int_0^\infty x^{((\nu+1)/2)-1} e^{-(\nu+t^2)x/2} dx \qquad \left(\begin{array}{c}\text{integrand is kernel of}\\ \text{gamma}((\nu{+}1)/2, 2/(\nu{+}t^2)\end{array}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \frac{\nu^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2}} \Gamma((\nu+1)/2) \left(\frac{2}{\nu+t^2}\right)^{(\nu+1)/2}$$

$$= \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)} \frac{1}{(1+t^2/\nu)^{(\nu+1)/2}},$$

the pdf of a t_ν distribution.

b. Differentiate both sides with respect to t to obtain

$$\nu f_F(\nu t) = \int_0^\infty y f_1(ty) f_\nu(y) dy,$$

where f_F is the F pdf. Now write out the two chi-squared pdfs and collect terms to get

$$\begin{aligned}
\nu f_F(\nu t) &= \frac{t^{-1/2}}{\Gamma(1/2)\Gamma(\nu/2)2^{(\nu+1)/2}} \int_0^\infty y^{(\nu-1)/2} e^{-(1+t)y/2} dy \\
&= \frac{t^{-1/2}}{\Gamma(1/2)\Gamma(\nu/2)2^{(\nu+1)/2}} \frac{\Gamma(\frac{\nu+1}{2})2^{(\nu+1)/2}}{(1+t)^{(\nu+1)/2}}.
\end{aligned}$$

Now define $y = \nu t$ to get

$$f_F(y) = \frac{\Gamma(\frac{\nu+1}{2})}{\nu\Gamma(1/2)\Gamma(\nu/2)} \frac{(y/\nu)^{-1/2}}{(1+y/\nu)^{(\nu+1)/2}},$$

the pdf of an $F_{1,\nu}$.

c. Again differentiate both sides with respect to t, write out the chi-squared pdfs, and collect terms to obtain

$$(\nu/m) f_F((\nu/m)t) = \frac{t^{-m/2}}{\Gamma(m/2)\Gamma(\nu/2)2^{(\nu+m)/2}} \int_0^\infty y^{(m+\nu-2)/2} e^{-(1+t)y/2} dy.$$

Now, as before, integrate the gamma kernel, collect terms, and define $y = (\nu/m)t$ to get

$$f_F(y) = \frac{\Gamma(\frac{\nu+m}{2})}{\Gamma(m/2)\Gamma(\nu/2)} \left(\frac{m}{\nu}\right)^{m/2} \frac{y^{m/2-1}}{(1+(m/\nu)y)^{(\nu+m)/2}},$$

the pdf of an $F_{m,\nu}$.

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

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

5.22 Calculating the cdf of Z^2, we obtain

$$\begin{aligned}
F_{Z^2}(z) &= P((\min(X,Y))^2 \le z) = P(-z \le \min(X,Y) \le \sqrt{z}) \\
&= P(\min(X,Y) \le \sqrt{z}) - P(\min(X,Y) \le -\sqrt{z}) \\
&= [1 - P(\min(X,Y) > \sqrt{z})] - [1 - P(\min(X,Y) > -\sqrt{z})] \\
&= P(\min(X,Y) > -\sqrt{z}) - P(\min(X,Y) > \sqrt{z}) \\
&= P(X > -\sqrt{z})P(Y > -\sqrt{z}) - P(X > \sqrt{z})P(Y > \sqrt{z}),
\end{aligned}$$

where we use the independence of X and Y. Since X and Y are identically distributed, $P(X > a) = P(Y > a) = 1 - F_X(a)$, so

$$F_{Z^2}(z) = (1 - F_X(-\sqrt{z}))^2 - (1 - F_X(\sqrt{z}))^2 = 1 - 2F_X(-\sqrt{z}),$$

since $1 - F_X(\sqrt{z}) = F_X(-\sqrt{z})$. Differentiating and substituting gives

$$f_{Z^2}(z) = \frac{d}{dz}F_{Z^2}(z) = f_X(-\sqrt{z})\frac{1}{\sqrt{z}} = \frac{1}{\sqrt{2\pi}}e^{-z/2}z^{-1/2},$$

the pdf of a χ^2_1 random variable. Alternatively,

$$\begin{aligned}
P(Z^2 \le z) &= P\left([\min(X,Y)]^2 \le z\right) \\
&= P(-\sqrt{z} \le \min(X,Y) \le \sqrt{z}) \\
&= P(-\sqrt{z} \le X \le \sqrt{z}, X \le Y) + P(-\sqrt{z} \le Y \le \sqrt{z}, Y \le X) \\
&= P(-\sqrt{z} \le X \le \sqrt{z} | X \le Y)P(X \le Y) \\
&\quad + P(-\sqrt{z} \le Y \le \sqrt{z} | Y \le X)P(Y \le X) \\
&= \frac{1}{2}P(-\sqrt{z} \le X \le \sqrt{z}) + \frac{1}{2}P(-\sqrt{z} \le Y \le \sqrt{z}),
\end{aligned}$$

using the facts that X and Y are independent, and $P(Y \le X) = P(X \le Y) = \frac{1}{2}$. Moreover, since X and Y are identically distributed

$$P(Z^2 \le z) = P(-\sqrt{z} \le X \le \sqrt{z})$$

and

$$\begin{aligned}
f_{Z^2}(z) &= \frac{d}{dz}P(-\sqrt{z} \le X \le \sqrt{z}) = \frac{1}{\sqrt{2\pi}}(e^{-z/2}\frac{1}{2}z^{-1/2} + e^{-z/2}\frac{1}{2}z^{-1/2}) \\
&= \frac{1}{\sqrt{2\pi}}z^{-1/2}e^{-z/2},
\end{aligned}$$

the pdf of a χ^2_1.

5.23

$$\begin{aligned}
P(Z > z) &= \sum_{x=1}^{\infty} P(Z > z|x)P(X = x) = \sum_{x=1}^{\infty} P(U_1 > z, \ldots, U_x > z|x)P(X = x) \\
&= \sum_{x=1}^{\infty}\prod_{i=1}^{x} P(U_i > z)P(X = x) \qquad \text{(by independence of the } U_i\text{'s)} \\
&= \sum_{x=1}^{\infty} P(U_i > z)^x P(X = x) = \sum_{x=1}^{\infty}(1 - z)^x \frac{1}{(e-1)x!} \\
&= \frac{1}{(e-1)}\sum_{x=1}^{\infty}\frac{(1-z)^x}{x!} = \frac{e^{1-z} - 1}{e - 1} \qquad 0 < z < 1.
\end{aligned}$$

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.

5.25 The joint pdf of $X_{(1)}, \ldots, X_{(n)}$ is

$$f(u_1,\ldots,u_n) = \frac{n!a^n}{\theta^{an}}u_1^{a-1}\cdots u_n^{a-1}, \quad 0<u_1<\cdots<u_n<\theta.$$

Make the one-to-one transformation to $Y_1 = X_{(1)}/X_{(2)}, \ldots, Y_{n-1} = X_{(n-1)}/X_{(n)}, Y_n = X_{(n)}$. The Jacobian is $J = y_2y_3^2\cdots y_n^{n-1}$. So the joint pdf of $Y_1,\ldots,Y_n$ is

$$\begin{aligned} f(y_1,\ldots,y_n) &= \frac{n!a^n}{\theta^{an}}(y_1\cdots y_n)^{a-1}(y_2\cdots y_n)^{a-1}\cdots(y_n)^{a-1}(y_2y_3^2\cdots y_n^{n-1}) \\ &= \frac{n!a^n}{\theta^{an}}y_1^{a-1}y_2^{2a-1}\cdots y_n^{na-1}, \quad 0<y_i<1; i=1,\ldots,n-1, \quad 0<y_n<\theta. \end{aligned}$$

We see that $f(y_1,\ldots,y_n)$ factors so $Y_1,\ldots,Y_n$ are mutually independent. To get the pdf of Y_1, integrate out the other variables and obtain that $f_{Y_1}(y_1) = c_1y_1^{a-1}$, $0<y_1<1$, for some constant c_1. To have this pdf integrate to 1, it must be that $c_1 = a$. Thus $f_{Y_1}(y_1) = ay_1^{a-1}$, $0<y_1<1$. Similarly, for $i = 2,\ldots,n-1$, we obtain $f_{Y_i}(y_i) = iay_i^{ia-1}, 0<y_i<1$. From Theorem 5.4.4, the pdf of Y_n is $f_{Y_n}(y_n) = \frac{na}{\theta^{na}}y_n^{na-1}$, $0<y_n<\theta$. It can be checked that the product of these marginal pdfs is the joint pdf given above.

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,

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