**Statistics 581, Problem Set 2 Solutions**

Wellner; 10/16/2002

**Reading:** Chapter 1, especially pages 13 - 17; start reading chapter 2; Ferguson pages 1-25.
**Due:** Wednesday, October 16, 2002.

1. Suppose that $Y$ is a random variable with $E(Y^2) < \infty$.
(a) Show that

$$Var(Y) = E\{Var(Y|X)\} + Var\{E(Y|X)\}\,;$$

i.e.

$$E(Y - EY)^2 = E\{E[(Y - E(Y|X))^2|X]\} + E\{[E(Y|X) - E(Y)]^2\}\,.$$

(b) Interpret (a) geometrically.
(c) Suppose that $Y \sim \chi^2_n(\delta)$. Compute $E(Y)$ and $Var(Y)$.
Hint: Use $E(Y) = E\{E(Y|X)\}$ and (a).
(d) Show that

$$\frac{\chi^2_n(\delta) - (n+\delta)}{\sqrt{2n+4\delta}} \rightarrow_d N(0,1)$$

as either $n \rightarrow \infty$ or $\delta \rightarrow \infty$.

**Solution:** (a) We compute directly:

$$\begin{aligned}
Var(Y) &= E[Y - E(Y)]^2 = E[Y - E(Y|X) + E(Y|X) - E(Y)]^2 \\
&= E[Y - E(Y|X)]^2 + 2E[(Y - E(Y|X)][E(Y|X) - E(Y)] \\
&\qquad + E[E(Y|X) - E(Y)]^2 \\
&= E\{E\{[Y - E(Y|X)]^2|X\}\} + 0 + Var[E(Y|X)] \\
&= E\{Var[Y|X]\} + Var[E(Y|X)]
\end{aligned}$$

since, by computing conditionally,

$$\begin{aligned}
&E[(Y - E(Y|X)][E(Y|X) - E(Y)] \\
&\quad = E\{E\{[(Y - E(Y|X)][E(Y|X) - E(Y)]|X\}\}
\end{aligned}$$

$$
\begin{aligned}
&= E\{[E(Y|X) - E(Y)]E\{[Y - E(Y|X)]|X\}\} \\
&= E\{[E(Y|X) - E(Y)]\{E(Y|X) - E(Y|X)\}\} \\
&= E\{[E(Y|X) - E(Y)] \cdot 0\} \\
&= 0\,.
\end{aligned}
$$

(b) A geometric interpretation of (a) is that $Y - E(Y|X)$ is orthogonal to $E(Y|X) - E(Y)$ in $L_2(\Omega, \mathcal{A}, P) = L_2(P)$, thus the identity in (a) can be intepreted as a "pythagorean theorem". Also note that $Y - E(Y|X)$ is orthogonal to any function $g(X)$: much as in the last part of (a)

$$
\begin{aligned}
&E[(Y - E(Y|X)]g(X) \\
&\quad = E\{E\{[(Y - E(Y|X)]g(X)|X\}\} \\
&\quad = E\{g(X)E\{[Y - E(Y|X)]|X\}\} \\
&\quad = E\{g(X)\{E(Y|X) - E(Y|X)\}\} \\
&\quad = E\{g(X) \cdot 0\} \\
&\quad = 0\,.
\end{aligned}
$$

(c) Now $(Y|K) \sim \chi^2_{2K+n}$ where $K \sim \text{Poisson}(\delta/2)$, so

$$
E(Y) = E\{E(Y|K)\} = E\{2K + n\} = n + 2(\delta/2) = n + \delta\,.
$$

Furthermore, using part (a) we get

$$
\begin{aligned}
Var(Y) &= E\{Var(Y|K)\} + Var\{E(Y|K)\} \\
&= E\{2(2K + n)\} + Var\{2K + n\} \\
&= 4(\delta/2) + 2n + 4(\delta/2) \\
&= 2n + 4\delta\,.
\end{aligned}
$$

(d) First note that if $Y \sim \chi^2_n(\delta)$, then $Y =_d (Z_1 + \sqrt{\delta})^2 + Z_2^2 + \cdots + Z_n^2$ where $Z_1, Z_2, \ldots, Z_n$ are independent, $Z_1, \ldots, Z_n \sim N(0,1)$. We can write, with $T_i \equiv Z_i^2 - 1$ for $i = 2, \ldots, n$ (having mean 0 and variance 2),

$$
\frac{\chi^2_n(\delta) - (n + \delta)}{\sqrt{2n + 4\delta}} \quad =_d \quad \frac{(Z_1 + \sqrt{\delta})^2 - (1 + \delta)) + (Z_2^1 - 1) + \cdots + (Z_n^2 - 1)}{\sqrt{2n + 4\delta}}
$$

$$
\begin{aligned}
&= \frac{(Z_1^2 - 1) + \cdots + (Z_n^2 - 1)}{\sqrt{2n + 4\delta}} + \frac{2\sqrt{\delta} Z_1}{\sqrt{2n + 4\delta}} \\
&= \frac{T_1 + \cdots + T_n}{\sqrt{2n}} \frac{\sqrt{2n}}{\sqrt{2n + 4\delta}} + \frac{2\sqrt{\delta} Z_1}{\sqrt{2n + 4\delta}}
\end{aligned}
$$

where, as $n \to \infty$,

$$
\frac{2\sqrt{\delta} Z_1}{\sqrt{2n + 4\delta}} \to_p 0; \qquad \text{and} \qquad \frac{T_1 + \cdots + T_n}{\sqrt{2n}} \to_d N(0,1)
$$

by the CLT. Hence by Slutsky's theorem

$$
\frac{\chi_n^2(\delta) - (n + \delta)}{\sqrt{2n + 4\delta}} \to_d 0 + N(0,1) \cdot 1 = N(0,1) \qquad \text{as} \quad n \to \infty \, .
$$

If $n$ is fixed and $\delta \to \infty$,

$$
\frac{T_1 + \cdots + T_n}{\sqrt{2n + 4\delta}} \to_p 0
$$

while

$$
\frac{2\sqrt{\delta} Z_1}{\sqrt{2n + 4\delta}} \to_d Z_1 \sim N(0,1) \, .
$$

Hence the desired conclusion follows from Slutsky's theorem (Proposition 2.2.9).

2. Suppose that: (i) $X \sim N_n(\mu, \Sigma)$ where $\Sigma$ is of rank $k < n$;
(ii) $\Sigma$ is a projection matrix (i.e. $\Sigma^2 = \Sigma$);
(iii) $\Sigma\mu = \mu$.
Show that $X'X \sim \chi_k^2(\delta)$ with $\delta = \mu'\mu$.

**Solution:** See Ferguson, ACILST, page 63.

3. (a) Ferguson, ACILST, #4, page 6:
Give an example of random variables $X_n$ such that $E|X_n| \to 0$ and $E|X_n|^2 \to 1$.
(b) Give an example of a sequence of random variables $X_n$ for which $X_n \to_p 0$ but $X_n \to_{a.s.} 0$ fails.

**Solution:** (a) If $X_n = a_n$ with probability $p_n$ and $X_n = 0$ with probability $1 - p_n$, then $E(X_n) = a_n p_n$ and $E(X_n^2) = a_n^2 p_n = 1$ if $p_n = 1/a_n^2$. Then $E(X_n) = a_n/a_n^2 = 1/a_n \to 0$ if $a_n \to \infty$. Ferguson's solution on page 173 takes $a_n = n$; the same holds for any sequence $a_n \to \infty$.
(b) Let $U \sim \text{Uniform}(0,1)$. The "dancing functions" are defined by $X_{n,k} = 1_{[(k-1)/2^n, k/2^n)}(U)$, $k = 1, \ldots, 2^n$, $n = 1, 2, \ldots$. Let $\{Y_m\}_{m \geq 1}$ be defined by $Y_m = X_{n,k}$ if $m = (\sum_{j=1}^{n} 2^j) + k = 2^{n+1} - 2 + k$ with $1 \leq k \leq 2^n$. Then for $\epsilon \in (0,1)$,

$$P(|Y_m| > \epsilon) = P(|X_{n,k}| > \epsilon) = 2^{-n} \to 0$$

so $Y_m \to_p 0$, but for every $U(\omega) \in (0,1)$ we have $Y_m(\omega) = 1$ for infinitely many $m$'s and also $Y_m(\omega) = 0$ for infinitely many $m$'s. Hence

$$0 = \liminf Y_m < \limsup Y_m = 1 \qquad a.s.$$

and it follows that $Y_m$ does not converge to 0 almost surely.

4. (a) If $W \sim \chi_2^2 = \text{Gamma}(2/2, 1/2) = \text{Gamma}(1, 1/2)$, find the density function $f_W$, distribution function $F_W$, and inverse distribution function $F_W^{-1}$ explicitly.
(b) Suppose that $(X, Y) \sim N_2(0, I)$. Show that $R$ and $\Theta$ defined by $R^2 = X^2 + Y^2$ and $\Theta = \arctan(Y/X)$ are independent random variables with $R^2 \sim \chi_2^2$ and $\Theta \sim \text{Uniform}(0, 2\pi)$.
(c) Use the results of (a) and (b) to show (using Theorem 2.3.4) how to use two independent Uniform$(0,1)$ random variables $U$ and $V$ to generate two standard normal random variables.

**Solution:** (a) If $W \sim \chi_2^2 = \text{Gamma}(1, 1/2)$, the density function is given by $f_W(w) = (1/2)e^{-w/2} 1_{[0,\infty)}$; i.e. $W \sim \text{Exponential}(1/2)$. Hence the distribution function is $F_W(w) = 1 - \exp(-w/2)$ for $w \geq 0$, and the inverse distribution function is $F_W^{-1}(u) = -2\log(1-u)$.
(b) The joint density of $(X, Y)$ is given by

$$f_{X,Y}(x,y) = \frac{1}{2\pi} \exp(-(x^2 + y^2)/2) \qquad \text{for} (x,y) \in R^2 .$$

Moreover, $X = R\cos(\Theta)$ and $Y = R\sin(\Theta)$. Hence the Jacobian of the transformation is

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \Big| \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} \Big| = r\cos^2(\theta) + r\sin^2(\theta) = r .$$

Thus we find that the joint density of $(R, \Theta)$ is given by

$$\begin{aligned} f_{R,\Theta}(r,\theta) &= f_{X,Y}(r\cos(\theta), r\sin(\theta)) = \frac{1}{2\pi}\exp(-r^2/2)r \qquad \text{on } (0,\infty)\times[0,2\pi] \\ &= r\exp(-r^2/2)\cdot\frac{1}{2\pi} = f_R(r)f_\Theta(\theta)\,. \end{aligned}$$

Thus $R$ and $\Theta$ are independent with densities $f_R(r) = r\exp(-r^2/2)1_{(0,\infty)}$ and $f_\Theta(\theta) = (2\pi)^{-1}1_{[0,2\pi]}(\theta)$. Note that the distribution function of $R$ is given by

$$F_R(r) = \int_0^r f_R(y)dy = \int_0^r y\exp(-y^2/2)dy = 1 - \exp(-r^2/2)\,.$$

It follows easily from this that

$$F_{R^2}(x) = P(R^2 \le x) = P(R \le \sqrt{x}) = 1 - \exp(-x/2)$$

for $x \in [0,\infty)$; i.e. $R^2 \sim$ Exponential$(1/2) =$ Gamma$(1, 1/2) = \chi_2^2$.
(c) If $U$ and $V$ are independent Uniform$(0,1)$ random variables, we can use the inverse transformation to first obtain

$$R^2 \equiv F_{\chi_2^2}^{-1}(U) = -2\log(1-U) \sim \chi_2^2 \qquad \text{and} \qquad \Theta \equiv 2\pi V \sim \text{Uniform}(0,2\pi)$$

note that $R^2$ and $\Theta$ are independent by independence of $U$ and $V$. Then in view of (b)

$$(X,Y) \equiv (R\cos(\Theta), R\sin(\Theta)) \sim N_2(0,I)\,.$$

5. Suppose that $U \sim$ Uniform$(0,1)$, $\alpha > 0$, and

$$X_n \equiv (n^\alpha/\log(n+1))1_{[0,1/n^\alpha]}(U)\,.$$

(a) Show that $X_n \to_{a.s.} 0$ and $E(X_n) \to E(0) = 0$.

(b) Can you find a random variable $Y$ with $|X_n| \le Y$ for all $n$ with $E(Y) < \infty$ for any $\alpha$?
(c) For what values of $\alpha$ does the uniform integrability condition

$$\limsup_{n\to\infty} E\{|X_n|1_{[|X_n|\ge M]}\} \to 0 \qquad \text{as } M \to \infty$$

hold?

**Solution:** (a) $X_n \to_{a.s.} 0$ since $X_n(\omega) = 0$ for $1/n^\alpha < U(\omega)$, or equivalently $n > (1/U(\omega))^{1/\alpha}$ and since $P(0 < U \le 1) = 1$. Moreover,

$$E(X_n) = \frac{n^\alpha}{\log(n+1)} \frac{1}{n^\alpha} = \frac{1}{\log(n+1)} \to 0 = E(0)\,.$$

(b) Now the smallest possible random variable $Y$ satisfying $|X_n| \le Y$ for all $n$ is $Y$ defined by

$$Y = \sum_{k=1}^\infty \frac{k^\alpha}{\log(k+1)} 1_{(1/(k+1)^\alpha, 1/k^\alpha]}(U)\,.$$

But we compute

$$\begin{aligned}
E(Y) &= \sum_{k=1}^\infty \frac{k^\alpha}{\log(k+1)} \left\{ \frac{1}{k^\alpha} - \frac{1}{(k+1)^\alpha} \right\} \\
&= \sum_{k=1}^\infty \frac{1}{\log(k+1)} \left\{ 1 - \left( \frac{k}{k+1} \right)^\alpha \right\} \\
&= \sum_{k=1}^\infty \frac{1}{\log(k+1)} \left\{ 1 - \left( 1 - \frac{1}{k+1} \right)^\alpha \right\} \\
&\ge \sum_{k=1}^{k(\alpha)} \frac{1}{\log(k+1)} \left\{ 1 - \left( 1 - \frac{1}{k+1} \right)^\alpha \right\} \\
&\qquad + \sum_{k=k(\alpha)}^\infty \frac{1}{\log(k+1)} \frac{\alpha/2}{k+1} \\
&= \infty\,.
\end{aligned}$$

since $(1-x)^\alpha \le 1 - \alpha x/2$ for $x \le x(\alpha)$. Thus there is no integrable dominating function $Y$ for any value of $\alpha$.
(c) On the other hand the uniform integrability condition does hold for any $\alpha > 0$:

$$\begin{aligned}
E\{|X_n| 1_{[|X_n| \ge M]}\} &= E\left\{ \frac{n^\alpha}{\log(n+1)} 1_{[0,1/n^\alpha]}(U) 1\{(n^\alpha/\log(n+1)) \ge M, U \le 1/n^\alpha\} \right\} \\
&= \frac{n^\alpha}{\log(n+1)} E\left\{ 1_{[0,1/n^\alpha]}(U) \right\} 1\{(n^\alpha/\log(n+1)) \ge M\}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\log(n+1)} 1\{(n^\alpha / \log(n+1)) \geq M\} \\
&\to 0 \cdot 1 = 0
\end{aligned}
$$

as $n \to \infty$ for every $\alpha > 0$ and $M > 0$. Hence the sequence $\{X_n\}$ is uniformly integrable for every $\alpha > 0$.