Statistics 581, Problem Set 2, Solutions
Wellner; 10/10/2008

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

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

(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 $K \sim$ 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

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

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

Solution: See Ferguson, ACILST, page 63 (and page 57). Find Γ orthogonal so that $\Gamma'\Sigma\Gamma = D$ where D is diagonal. Now $\Gamma\Gamma' = I$, so if $\Sigma^2 = \Sigma$ we have $D^2 = \Gamma'\Sigma\Gamma\Gamma'\Sigma\Gamma = \Gamma'\Sigma^2\Gamma = \Gamma'\Sigma\Gamma = D$ and conversely. Moreover, since Σ is of rank k, D is of rank k, and this together with $D^2 = D$ implies that D has k 1's on the diagonal and $n-k$ 0's. Without loss, assume that Γ has been chosen so that the k ones occur in the the first r positions of the diagonal matrix D; thus

$$D = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

where I is $k \times k$. Moreover, note that

$$\begin{aligned} D\Gamma'\mu &= \Gamma'\Sigma\Gamma\Gamma'\mu = \Gamma'\Sigma\mu \\ &= \Gamma\mu, \end{aligned}$$

and this implies that the last $n-k$ components of $\Gamma'\mu$ are all zero. Now let $Y = \Gamma'X$ (much as in the proof of theorem 1.3.2 of the notes). Then $Y \sim N_n(\Gamma'\mu, D)$, $Y'Y = X'\Gamma\Gamma'X = X'X$, and by (d) of page 16, section 1.3, $X'X = Y'Y \sim \chi^2_k(\delta)$ where $\delta = (\Gamma'\mu)'(\Gamma'\mu) = \mu'\Gamma\Gamma'\mu$.

3. Ferguson, ACILST, #1, page 11.
Let $X_1, X_2, \ldots$ be i.i.d. random variables with densities $f(x) = \alpha x^{-(\alpha+1)} 1_{(1,\infty)}(x)$.
(a) For what values of $\alpha > 0$ and $r > 0$ is it true that $n^{-1}X_n \to_r 0$?
(b) For what values of $\alpha > 0$ is it true that $n^{-1}X_n \to_{a.s.} 0$?
(c) If $X_1, X_2, \ldots$ are independent with X_n having density $f_n(x) = \alpha_n x^{-(\alpha_n+1)} 1_{(1,\infty)}(x)$ for $n = 1, 2, \ldots$, Find the limit of $n^{-2}EX_n^2$ when $\alpha_n = 2 + n^{-\gamma}$ for $\gamma \in \mathbb{R}$.

Solution: (a) $E(X_n^r) = \alpha \int_1^\infty x^r x^{-(\alpha+1)} dx = \alpha/(\alpha - r)$ if $\alpha > r$, while $E(X_n^r) = \infty$ if $\alpha \le r$. Thus $E[(n^{-1}X_n)^r] = n^{-r}(\alpha/(\alpha - r)) \to 0$ if $\alpha > r$.
(b) $P(n^{-1}X_n > \epsilon) = P(X_n > n\epsilon) = \alpha \int_{n\epsilon}^\infty x^{-(\alpha+1)} dx = (n\epsilon)^{-\alpha}$, so for $\alpha > 1$ we have

$$\sum_{n=1}^\infty P(n^{-1}X_n > \epsilon) \le \sum_{n=1}^\infty (n\epsilon)^{-\alpha} < \infty.$$

Thus $P(n^{-1}X_n > \epsilon \text{ i.o.}) = 0$ by the Borel-Cantelli lemma, and we conclude that $n^{-1}X_n \to_{a.s.} 0$ for $\alpha > 1$. [Since the X_n's are independent, $n^{-1}X_n \to_{a.s.} 0$ if *and only if* $\alpha > 1$ by the converse Borel-Cantelli lemma.]
(c) If the X_n's are independent with the same densities as in (a) and (b) but with $\alpha = \alpha_n = 2 + n^{-\gamma}$ for X_n, then $EX_n^2 = (2 + n^{-\gamma})/n^{-\gamma}$, so

$$n^{-2}EX_n^2 = (2 + n^{-\gamma})/n^{2-\gamma} \to \begin{cases} 0, & \text{if } \gamma < 2, \\ 2, & \text{if } \gamma = 2, \\ \infty, & \text{if } \gamma > 2. \end{cases}$$

4. (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 random variables X_n such that $E|X_n| \to 0$ and $E|X_n|^2 \to \infty$.
(c) 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 $U \sim$ Uniform$(0, 1)$, and set $X_n = n^{\alpha} 1_{[0,1/n]}(U)$. Then $EX_n = n^{\alpha} n^{-1} \to 0$ if $\alpha < 1$, while $EX_n^2 = n^{2\alpha} n^{-1} \to \infty$ if $\alpha > 1/2$. Thus the required convergences hold for all $1/2 < \alpha < 1$.
(c) Let $U \sim$ 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 $\{Y_m\}_{m \ge 1}$ be defined by $Y_m = X_{n,k}$ if $m = (\sum_{j=1}^{n} 2^j) + k = 2^{n+1} - 2 + k$ with $1 \le k \le 2^n$. 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.

5. Continuation of problem 3, Problem set 1: Suppose that $X \sim F$ on $R^+ \equiv [0, \infty)$, $Y \sim G$ on R^+, and X and Y are independent random variables, but that only Y and $\Delta = 1\{X \le Y\}$ are observed. (This is called "type 1 interval censored data", or "current status data".)
(a) Find the joint distribution of (Y, Δ) by first computing the distribution of Δ conditional on Y, and then computing the two subdistribution functions $F_1(y) \equiv P(Y \le y, \Delta = 1)$ and $F_0(t) \equiv P(Y \le y, \Delta = 0)$.
(b) If $X \sim$ Exponential(λ) and $Y \sim$ Exponential(μ), are Y and Δ independent?

Solution: (a) When Z is replaced by $\tilde{Z} = Y$, then since $\tilde{Z} = Y$ and $\Delta = 1\{X \le Y\}$, it follows that $(\Delta | Y) \sim$ Bernoulli$(F(Y))$. Thus

$$p_F(\delta | y) = F(y)^{\delta} (1 - F(y))^{1-\delta}, \qquad \delta \in \{0, 1\},$$

and this yields

$$F_1(y) \equiv P(Y \le y, \Delta = 1) = P(Y \le y, X \le Y) = \int_{[0,y]} F(z) dG(z)\,,$$

and

$$F_0(y) \equiv P(Y \le y, \Delta = 0) = P(Y \le y, X > Y) = \int_{[0,y]} (1 - F(z)) dG(z)\,.$$

(b) When $X \sim$ Exponential(λ) and $Y \sim$ Exponential(μ) the marginal distributions are given by

$$P(Y \le y) = 1 - \exp(-\mu y),$$
$$P(\Delta = 1) = P(X \le Y) = \int_0^\infty e^{-\mu x} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda + \mu},$$

but now we have

$$\begin{aligned} F_1(y) &= P(Y \le y, \Delta = 1) = P(Y \le y, X \le Y) \\ &= \int_0^y F(z) dG(z) = \int_0^y (1 - \exp(-\lambda z)) \mu e^{-\mu z} dz \\ &= (1 - e^{-\mu y} - \frac{\mu}{\lambda + \mu}(1 - \exp(-(\lambda + \mu) y)) \\ &\ne P(\Delta = 1) P(Y \le y) \end{aligned}$$

so Y and Δ are *not* independent for exponential X and Y.

6. (a) Lehmann and Casella, #3.5, page 64.
   (b) Lehmann and Casella, #3.6, page 64.
   (c) Lehmann and Casella, #3.7, page 64.

Solution: (a) (i) Suppose that S is not closed. Then there exists a sequence $\{x_n\} \subset S$ such that $x_n \to x_0 \in S^c$. But then, for every $\epsilon > 0$ there is an open ball $B(x_0, \epsilon)$ such that $x_n \in B(x_0, \epsilon)$ for $n \ge N_\epsilon$. Since each x_n is a support point, $P(B(x_0, \epsilon)) > 0$ for each $\epsilon > 0$. But for any open set A with $x_0 \in A$, $B(x_0, \epsilon) \subset A$ for some $\epsilon > 0$, and hence $P(A) \ge P(B(x_0, \epsilon) > 0$. But this implies $x_0 \in S$. Contradiction. Thus S is closed.
(ii) $P(S) = 1$. From (i) S is closed, so S^c is open. Since $x \in S^c$ if and only if $x \in A_x$ for some open rectangle A_x satisfying $P(A_x) = 0$. Thus $S^c \subset \cup_x A_x$. By the Lindelöf theorem (see e.g. Apostol (1964), page 51), for any such open covering $\{A_x\}_{x \in S^c}$ of $S^c \subset R^d$, there is a countable subcollection $\{A_{x_n}\}$ which covers S^c: $S^c \subset \cup_n A_{x_n}$. Then we have

$$P(S^c) \le P(\cup_n A_{x_n}) \le \sum_n P(A_{x_n}) = \sum_n 0 = 0\,.$$

Hence $P(S) = 1$.
(iii) We want to show that $S = \cap\{C : \ C \text{ closed}, P(C) = 1\}$. From

(i) and (ii) we know that S is in the collection of sets on the right side, so it follows that $S \supset \cap\{C : \; C \text{ closed}, P(C) = 1\}$. Thus it remains to show that $S \subset \cap\{C : \; C \text{ closed}, P(C) = 1\}$. Equivalently, it remains to show that $S^c \supset \cup\{C^c : C^c \text{ open}, \; P(C^c) = 0\}$. But if $x \in \cup\{C^c : C^c \text{ open}, \; P(C^c) = 0\}$, then $x \in C^c$ for some C^c open with $P(C^c) = 0$, and hence also $x \in A \subset C^c$ for some open rectangle A (an open ball centered at x for the metric $\|y\| = \max_{1 \le i \le d} |x_i|$) with $P(A) \le P(C^c) = 0$. Hence $x \in S^c$.
(b) Suppose that P and Q are equivalent: i.e. $Q \prec\prec P$ and $P \prec\prec Q$. Then for any open set A, $P(A) = 0$ if and only if $Q(A) = 0$. This implies that for any closed set A^c,

$$P(A^c) = 1 \qquad \text{if and only if} \qquad Q(A^c) = 1\,.$$

This implies that the minimal closed set S_P with $P(S_P) = 1$ is also the minimal closed set S_Q with $Q(S_Q) = 1$; i.e. $S_P = \text{supp}(P) = \text{supp}(Q) = S_Q$.
(c) Since $P(X = 1/n) = p_n > 0$ for $n = 1, 2, \ldots$ with $\sum_1^\infty p_n = 1$, it follows that $\text{supp}(P) = \{0, \ldots, 1/n, \ldots, 1/2, 1\}$, which is closed. Similarly, Since $Q(X = 1/n) = q_n > 0$ for $n = 1, 2, \ldots$ with $\sum_1^\infty q_n = 1/2$, and $Q(X = 0) = 1/2$, it follows that $\text{supp}(Q) = \{0, \ldots, 1/n, \ldots, 1/2, 1\} = \text{supp}(P)$. But $P(\{0\}) = 0$ while $Q(\{0\}) = 1/2$, so $Q \prec\prec P$ fails. Thus Q and P are not equivalent.