Statistics 581, Problem Set 1 Solutions

Wellner; 10/05/2005

1. Let X and Y be i.i.d. Uniform$(0,1)$ random variables Define $U = X+Y$, $V = \min(X,Y) = X \wedge Y$.

   (i) What is the range of (U,V)?

   (ii) Find the joint density function $f_{U,V}(u,v)$ of the pair (U,V). Are U and V independent?

Solution: (i) The range of (X,Y) is
$A = \{(x,y) : 0 \le x \le 1,\ 0 \le y \le 1\}$. The range of (U,V) is

$$\begin{aligned} B &= \{(u,v) : 0 \le v \le 1, 2v \le u \le 1+u\} \\ &= \{(u,v) :\ 0 \le u \le 1,\ 0 \le v \le u/2\} \\ &\qquad \cup \{(u,v) :\ 1 < u \le 2,\ u-1 \le v \le v/2\}. \end{aligned}$$

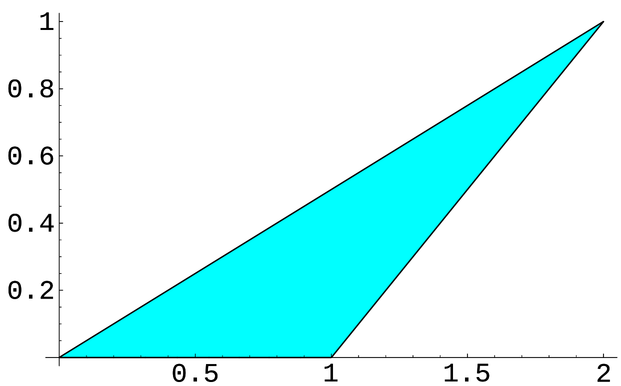

Figure 1: Range of U,V .

(ii) First solution - via Jacobians: The transformation $(X,Y) \to (U,V)$ is 1-1 and onto from A to B. On the set $x < y$, its inverse is given by $X = V$, $Y = U - V$; on the set $x > y$, its inverse is given by

$X = U - V$, $Y = V$. These mappings are continuously differentiable on $B^* \equiv B \setminus \{(u,v): \ v = u/2\} = B \setminus$ a null set. On B^* the Jacobian of the transformations are

$$\det \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = -1 \qquad \text{if } x < y, \qquad \det \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = 1 \qquad \text{if } x > y. \tag{1}$$

Thus by the usual transformation of densities formula, the joint density of (U,V) is obtained from $f_{X,Y}(x,y) = 1_{[0,1]}(x)1_{[0,1]}(y)$ as follows:

$$\begin{aligned} f_{U,V}(u,v) &= f_{X,Y}(x(u,v), y(u,v))|\det \frac{\partial(x,y)}{\partial(u,v)}|1_{[0<x(u,v)<y(u,v)<1]} \\ &\quad + f_{X,Y}(x(u,v), y(u,v))|\det \frac{\partial(x,y)}{\partial(u,v)}|1_{[1>x(u,v)>y(u,v)>0]} \\ &= \left(1_{[0,1]}(v)1_{[0,1]}(u-v)1_{[0<v<u-v<1]} + 1_{[0,1]}(u-v)1_{[0,1]}(v)1_{[1>u-v>v>0]}\right) \\ &= 2 \cdot 1_{B^*}(u,v)\,. \end{aligned}$$

Thus the joint density of (U,V) is uniform on B^* (or uniform on B). The random variables U and V are clearly *not* independent since the range of (U,V) is not a product set in R^2; moreover, the joint density of (U,V) does not factor into the product of its marginal densities. The marginal densities are given by

$$f_V(v) = \int f_{U,V}(u,v)du = \int_{2u}^{1+u} 2du = 2(1-u), \quad u \in [0,1]$$

and

$$f_U(u) = \int f_{U,V}(u,v)du = \begin{cases} \int_0^{u/2} 2dv = u & 0 \le u \le 1 \\ \int_{u-1}^{u/2} 2dv = 2-u & 1 < u \le 2. \end{cases}$$

Second solution by direction calculation of the joint distribution function: Note that we can write

$$\begin{aligned} &P(U \le u, V > v) \\ &\quad = P(X+Y \le u, X \wedge Y > v) = P(X+Y \le u, X > v, Y > v) \\ &\quad = \begin{cases} \frac{1}{2}(u-2v)^2, & \text{if } \ 2v \le u \le 1+v, \\ (1-v)^2 - \frac{1}{2}(2-u)^2, & \text{if } \ u > 1+v\,. \end{cases} \end{aligned}$$

(This is easy by pictures!) Since

$$F_{U,V}(u,v) = P(U \le u, V \le v) = P(U \le u) - P(U \le u, V > v)$$

computing $(\partial^2/\partial u \partial v)P(U \le u, V \ge v)$ on each of these pieces separately again yields $f_{U,V}(u,v) = 2 \cdot 1_B(u,v)$. Also note that the marginal distribution functions of U and V are given by $F_U(u) = (1/2)u^2 1_{[0,1)}(u) + \{1 - \frac{1}{2}(2-u)^2\} 1_{[1,2]}(u)$ on $0 \le u \le 2$ and $F_V(v) = 1 - (1-v)^2$ for $0 \le v \le 1$.

2. Ferguson, ACILST, #2, page 6.
Solution: If $X_n \sim$ Uniformly on $\{1/n, 2/n, \ldots, 1\}$ then

$$F_n(t) \quad = \quad P(X_n \le t) = \frac{\#\text{of } j/n \le t}{n} = \frac{\lfloor nt \rfloor}{n}.$$

Note that $|F_n(t) - t| \le 1/n$ for each fixed $0 \le t \le 1$. Then $F_n(t) \to F(t) = t$ for all $0 \le t \le 1$. That is, $X_n \to_d X \sim \text{Uniform}(0,1)$. These X_n's do not necessarily converge in probability to X because the random variables X_n are not necessarily defined on the same probability space.

3. (Continuation of the previous problem). Now suppose that $U \sim \text{Uniform}(0,1)$ and for each $n \ge 1$ define $V_n \equiv \sum_{j=1}^n (j/n) 1_{((j-1)/n, j/n]}(U)$.
(a) Show that $V_n \stackrel{d}{=} X_n$ where X_n is as in problem 2.
(b) Show that $V_n \to_p U$.

Solution: (a) Note that $P(V_n = j/n) = P(U \in ((j-1)/n, j/n]) = 1/n = P(X_n = j/n)$. Thus $V_n \stackrel{d}{=} X_n$.
(b) To see that $V_n \to_p U$, note that

$$P(|V_n - U| > \epsilon) = \begin{cases} n(1/n - \epsilon) = 1 - n\epsilon\,, & \text{if } 1/n > \epsilon \\ 0\,, & \text{if } 1/n \le \epsilon\,. \end{cases}$$

Hence it follows that $V_n \to_p U$.

4. Ferguson, ACILST, #6, page 7. (This is known as the Polya-Cantelli lemma; see Chapter 2, Propostion 2.11, page 10.) If F is continuous, then $F_n \to_d F$ means that $F_n(x) \to F(x)$ for all $x \in R$. Show that in this case $\sup_x |F_n(x) - F(x)| \to 0$.
Solution. See Ferguson, ACILST page 173.

5. Suppose that F is the distribution function of random variables X, Y with $X \sim \text{Uniform}(0,1)$ marginally and $Y \sim \text{Uniform}(0,1)$ marginally. Thus $F(x,y) = P(X \le x, Y \le y)$ satisfies

$$F(x,1) = x, \quad 0 \le x \le 1, \qquad \text{and} \qquad F(1,y) = y, \quad 0 \le y \le 1.$$

(a) Show that

$$F(x,y) \le x \wedge y \equiv F_U(x,y)$$

for all $0 \le x \le 1$, $0 \le y \le 1$. Here $x \wedge y \equiv x$ if $x \le y$, y if $y \le x$.
(b) Show that

$$F(x,y) \ge (x+y-1)^+ \equiv F_L(x,y)$$

for all $0 \le x \le 1$, $0 \le y \le 1$. Here $z^+ = z 1_{[0,\infty)}(z)$.
(c) Show that F_U is the distribution function of (X,X) where $X \sim \text{Uniform}(0,1)$. Show that F_L is the distribution function of $(X, 1-X)$ where $X \sim \text{Uniform}(0,1)$.
(d) The distribution functions F_U and F_L are called the Fréchet bounds. Show that F_L and F_U are singular with respect to Lebesgue measure λ_2 on $[0,1]^2$; i.e. show that the corresponding probability measures P_L and P_U satisfy

$$P((X,Y) \in A) = 1, \qquad \lambda_2(A) = 0$$

and

$$P((X,Y) \in A^c) = 0, \qquad \lambda_2(A^c) = 1$$

for some set A (which will be different for P_L and P_U). This implies that F_L and F_U do not have densities with respect to Lebesgue measure on $[0,1]^2$. (See Chapter 0, Section 3, especially Definition 3.1 and Theorem 3.1.)

Solution: (a) By monotonicity of F in each argument,

$$F(x,y) \le F(x,1) = x \qquad \text{and} \qquad F(x,y) \le F(1,y) = y$$

for $0 \le x, y \le 1$. Hence it follows that

$$F(x,y) \le \min\{x,y\} \equiv F_U(x,y), \qquad 0 \le x, y \le 1.$$

(b) Note that

$$0 \le P(X > x, Y > y) = 1 - F(x,1) - F(1,y) + F(x,y) = 1 - x - y + F(x,y)\,,$$

and hence

$$F(x,y) \ge x + y - 1, \qquad 0 \le x, y \le 1\,.$$

Since $F(x,y) = P(X \le x, Y \le y) \ge 0$ trivially, it follows that

$$F(x,y) \ge (x+y-1) \vee 0 = (x+y-1)^+ \equiv F_L(x,y) \qquad 0 \le x, y \le 1\,.$$

(c) If $X \sim \text{Uniform}(0,1)$, the joint distribution function of the pair (X,X) is given by

$$P(X \le x, X \le y) = P(X \le x \wedge y) = x \wedge y$$

for $0 \le x, y \le 1$. Thus F_U is the distribution function of (X,X). Similarly, the joint distribution function of the pair $(X, 1-X)$ is given by

$$\begin{aligned} P(X \le x, 1 - X \le y) &= P(X \le x, X \ge 1-y) \\ &= \begin{cases} P(1-y \le X \le x) & \text{if } x \ge 1-y \\ 0 & \text{if } x < 1-y \end{cases} \\ &= \begin{cases} x - (1-y) & \text{if } x \ge 1-y \\ 0 & \text{if } x < 1-y \end{cases} \\ &= (x+y-1)^+\,. \end{aligned}$$

Thus F_L is the distribution function of $(X, 1-X)$.
(d) It is clear from part (c) that if we take $A = \{(x,x) : \ 0 \le x \le 1\} \subset [0,1]^2$, then $P_U((X,Y) \in A) = 1$ and $\lambda_2(A) = 0$. Similarly, if $B = \{(x, 1-x) : \ 0 \le x \le 1\} \subset [0,1]^2$, then $P_L((X,Y) \in B) = 1$ and $\lambda_2(B) = 0$. Thus both P_U and P_L are singular with respect to Lebesgue measure λ_2 on $[0,1]^2$, and the hypothesis of the Radon-Nikodym theorem 0.3.1 is not satisfied. Hence neither P_U nor P_L have densities with respect to Lebesgue measure λ_2. Also note that

$$\frac{\partial^2}{\partial x \partial y} F_U(x,y) = 0 \ \text{ for } \ (x,y) \notin A\,,$$

while

$$\frac{\partial^2}{\partial x \partial y} F_L(x,y) = 0 \ \text{ for } \ (x,y) \notin B\,.$$

[Note that the arguments in (a) and (b) extend to an arbitrary distribution function F on R^2 with marginal d.f.'s F_X and F_Y respectively:

$$F(x,y) \le F_X(x) \wedge F_Y(y)\,,$$

and

$$F(x,y) \ge (F_X(x) + F_Y(y) - 1)^+\,.]$$

6. (a) Lehmann and Casella, TPE, problem 1.2, page 62.
(b) Lehmann and Casella, TPE, problem 1.3, page 62.

Solution: Solution: (i) First solution – via the Cauchy-Schwarz inequality: First recall the Cauchy-Schwarz inequality in R^n: if $u, v \in R^n$, then $(u'v)^2 \le (u'u)(v'v)$ with equality iff $u = cv$ for some real number c. Now extend this as follows: if Σ is positive definite and $x, y \in R^n$, then

$$(x'y)^2 \quad = \quad (\Sigma^{1/2}x)'(\Sigma^{-1/2}y) \le (x'\Sigma x)(y'\Sigma^{-1}y)$$

with equality iff $\Sigma^{1/2}x = c\Sigma^{-1/2}y$; i.e. iff $x = c\Sigma^{-1}y$.

Now consider X, a random vector in R^n, with $E(X) = \mathbf{1}\theta$ and $Cov(X,X) = E[(X - E(X))(X - E(X))'] = \Sigma$, where $\mathbf{1} = (1, \ldots, 1)' \in R^n$. A linear estimator $\alpha'X = \alpha_1 X_1 + \ldots + \alpha_n X_n$ is unbiased for θ iff $\theta = E(\alpha'X) = \alpha'E(X) = (\alpha'\mathbf{1})\theta$ for all θ; i.e., iff $\alpha'\mathbf{1} = 1$. The variance of $\alpha'X$ is $Var(\alpha'X) = \alpha'\Sigma\alpha$. To find the best such estimator, we must find

$$\min\{\alpha'\Sigma\alpha : \alpha'\mathbf{1} = 1\}.$$

But by the Cauchy-Schwarz inequality, if $\alpha'\mathbf{1} = 1$, then

$$\alpha'\Sigma\alpha \ge 1/(\mathbf{1}'\Sigma^{-1}\mathbf{1})$$

with equality iff $\alpha = c\Sigma^{-1}\mathbf{1}$. The condition $\alpha'\mathbf{1} = \mathbf{1}$ then implies that $c = 1/(\mathbf{1}'\Sigma^{-1}\mathbf{1})$, so the optimal α is $\alpha_0 \equiv \Sigma^{-1}\mathbf{1}/(\mathbf{1}'\Sigma^{-1}\mathbf{1})$, and the optimal linear unbiased estmator is $\alpha_0'X = (\mathbf{1}'\Sigma^{-1}X)/(\mathbf{1}'\Sigma^{-1}\mathbf{1})$ whose variance is $Var(\alpha_0'X) = 1/(\mathbf{1}'\Sigma^{-1}\mathbf{1})$.
The solutions to 1.2, 1.3(a), and 1.3(b) now follow:
1.2: In this case $\Sigma = \sigma^2 I$, so $\alpha_0 = \mathbf{1}(1/\sigma^2)/(\mathbf{1}'I\mathbf{1}/\sigma^2) = \mathbf{1}(1/n)$.
1.3(a): The inverse of the matrix $\text{diag}(1/c_i)$ is just $\text{diag}(c_i)$. This

implies that $\alpha_0' X = (\sum_1^n a_i X_i)/(\sum_1^n c_i)$ and $Var(\alpha_0' X) = \sigma^2 / \sum c_i$.
1.3(b): The inverse of the matrix with 1 on the diagonal and ρ off the diagonal is of the form a in the diagonal entries and b in the off-diagonal entries for some a, b. Hence $\Sigma^{-1}\mathbf{1} = \sigma^{-2}(a + (n-1)b)\mathbf{1}$, which leads to $\mathbf{1}'\Sigma^{-1}X = \sigma^{-2}(a + (n-1)b)(X_1 + \ldots + X_n)$, and $\mathbf{1}'\Sigma^{-1}\mathbf{1} = \sigma^{-2}(a + (n-1)b)n$. Hence we find that $\alpha_0' X = \sum_1^n X_i/n$. But $\Sigma\mathbf{1} = \sigma^2(1 + (n-1)\rho)\mathbf{1}$, so $\mathbf{1} = \sigma^2(1 + (n-1)\rho)\Sigma^{-1}\mathbf{1} = (1 + (n-1)\rho)(a + (n-1)b)\mathbf{1}$, and hence $[a + (n-1)b] = [1 + (n-1)\rho]^{-1}$. Therefore

$$Var(\alpha_0' X) = \frac{\sigma^2}{n}[1 + (n-1)\rho] \begin{cases} > \sigma^2/n & \text{if } \rho > 0 \\ < \sigma^2/n & \text{if } -1/(n-1) \le \rho < 0 \end{cases} .$$

[Note that if $\rho < -1/(n-1)$, the matrix Σ of this form *is not a covariance matrix*!