

Statistics 521, Problem Set 8 Solutions

Wellner; 11/21/2012

1. PfS Exercise 6.4.3, page 131. Prove just the parts of these formulas involving F , not the parts involving F^{-1} .

Solution: (11): Since $X \geq 0$ we may write $X = \int_0^X dt$ to obtain

$$\begin{aligned}
 E(X) &= \int_{\Omega} X dP = \int_{\Omega} \int_0^X dt dP \\
 &= \int_{\Omega} \int_0^{\infty} 1_{\{0 \leq t < X\}} dt dP \\
 &= \int_0^{\infty} \int_{\Omega} 1_{\{0 \leq t < X\}} dP dt \quad \text{by Fubini's theorem} \\
 &= \int_0^{\infty} P(X > t) dt = \int_0^{\infty} (1 - F(t)) dt.
 \end{aligned}$$

If either side is $+\infty$, the both sides are $+\infty$. (12): Since $X = X^+ - X^-$ with $X^+, X^- \geq 0$, it follows from (11) applied to X^+ and X^- (and assuming that $E|X| = EX^+ + EX^- < \infty$) that

$$\begin{aligned}
 E(X) &= E(X^+) - E(X^-) \\
 &= \int_0^{\infty} P(X^+ > t) dt - \int_0^{\infty} P(X^- > t) dt \\
 &= \int_0^{\infty} P(X > t) dt - \int_0^{\infty} P(-X > t) dt \\
 &= \int_0^{\infty} (1 - F(t)) dt - \int_0^{\infty} P(X < -t) dt \\
 &= \int_0^{\infty} (1 - F(t)) dt - \int_0^{\infty} F(-t-) dt \\
 &= \int_0^{\infty} (1 - F(t)) dt - \int_{-\infty}^0 F(t-) dt \\
 &= \int_0^{\infty} (1 - F(t)) dt - \int_{-\infty}^0 F(t) dt
 \end{aligned}$$

where the last equality follows by a change of variables and the fact that $F(t)$ differs from $F(t-)$ on a set which has Lebesgue measure at most zero.

(13): Since $X \geq 0$ we may write $X^r = \int_0^X rt^{r-1}dt$ to obtain

$$\begin{aligned} E(X^r) &= \int_{\Omega} X dP = \int_{\Omega} \int_0^X rt^{r-1} dt dP = \int_{\Omega} \int_0^{\infty} 1_{\{0 \leq t < X\}} rt^{r-1} dt dP \\ &= \int_0^{\infty} \int_{\Omega} 1_{\{0 \leq t < X\}} dP rt^{r-1} dt \quad \text{by Fubini's theorem} \\ &= \int_0^{\infty} rt^{r-1} P(X > t) dt = \int_0^{\infty} rt^{r-1} (1 - F(t)) dt. \end{aligned}$$

If either side is $+\infty$, then both sides are $+\infty$. (14): First suppose that X and Y are both non-negative and that G and H satisfy $G_-(0) = G_+(0) = H_-(0) = H_+(0) = 0$. Then we can write $G(x) = \int_{[0,x)} 1_{[0,x)}(s) dG_-(s)$, $H(y) = \int_{[0,y)} 1_{[0,y)}(t) dH_-(t)$. Now finiteness of the covariance $Cov[G(X), H(Y)]$ implies that $(G(X) - EG(X))(H(Y) - EH(Y))$ is integrable and hence that $G(X)H(Y)$ is integrable. Furthermore

$$\begin{aligned} Cov[G(X), H(Y)] &= E\{[G(X) - EG(X)][H(Y) - EH(Y)]\} \\ &= E\{G(X)H(Y)\} - EG(X) \cdot EH(Y). \end{aligned}$$

Thus Fubini's theorem yields

$$\begin{aligned} E\{G(X)H(Y)\} &= \int_{[0,\infty)} \int_{[0,\infty)} G(x)H(y) dF(x, y) \\ &= \int_{[0,\infty)} \int_{[0,\infty)} \left\{ \int_{[0,x)} 1_{[0,x)}(s) dG_-(s) \int_{[0,y)} 1_{[0,y)}(t) dH_-(t) \right\} dF(x, y) \\ &= \int_{[0,\infty)} \int_{[0,\infty)} \left\{ \int_{[0,x)} \int_{[0,y)} 1_{[0,x)}(s) 1_{[0,y)}(t) dF(x, y) \right\} dG(s) dH(t) \\ &\quad \text{by Fubini's theorem} \\ &= \int_{[0,\infty)} \int_{[0,\infty)} P(X > s, Y > t) dG(s) dH(t); \end{aligned}$$

note that this is always finite. Furthermore

$$\begin{aligned}
 EG(X) &= \int_{[0,\infty)} G(x)dF_X(x) = \int_{[0,\infty)} \int_{[0,\infty)} 1_{[0,x)}(s)dG_-(s)dF_X(x) \\
 &= \int_{[0,\infty)} \left\{ \int_{[0,\infty)} 1_{[0,x)}(s)dF_X(x) \right\} dG(s) \\
 &= \int_{[0,\infty)} P(X > s)dG(s) < \infty
 \end{aligned}$$

and, similarly,

$$EH(Y) = \int_{[0,\infty)} P(Y > t)dH(t) < \infty.$$

Combining these yields

$$\begin{aligned}
 E\{G(X)H(Y)\} - EG(X) \cdot EH(Y) \\
 &= \int_0^\infty \int_0^\infty \{P(X > s, Y > t) - P(X > s)P(Y > t)\}dG(s)dH(t) \\
 &= \int_0^\infty \int_0^\infty \{F(s, t) - F_X(s)F_Y(t)\}dG(s)dH(t)
 \end{aligned}$$

using the relations

$$\begin{aligned}
 P(X > s, Y > t) &= 1 - F_X(s) - F_Y(t) + F(s, t), \\
 P(X > s)P(Y > t) &= (1 - F_X(s))(1 - F_Y(t)) \\
 &= 1 - F_X(s) - F_Y(t) + F_X(s)F_Y(t).
 \end{aligned}$$

For general X, Y we can repeat this argument for each of the four quadrants, with each quadrant contributing a finite term, and add the results to obtain the claimed formula.

2. Prove the two formulas in (17), PfS page 131.

Solution: First note that if $X \geq 0$ is integer-valued, then the distribution function F of X is constant between integers, and $P(X > x) = 1 - F(x) = P(X > k)$ for $k \leq x < k + 1$. Thus from (7.4.11) we find

that

$$\begin{aligned}
 E(X) &= \int_0^{\infty} (1 - F(x)) dx \\
 &= \sum_{k=0}^{\infty} \int_{[k, k+1)} (1 - F(x)) dx \\
 &= \sum_{k=0}^{\infty} (1 - F(k)) \int_{[k, k+1)} dx \\
 &= \sum_{k=0}^{\infty} P(X > k) = \sum_{k=0}^{\infty} P(X \geq k + 1) \\
 &= \sum_{m=1}^{\infty} P(X \geq m).
 \end{aligned}$$

Similarly, from (7.4.13) with $r = 2$,

$$\begin{aligned}
 E(X^2) &= \int_0^{\infty} 2x(1 - F(x)) dx \\
 &= \sum_{k=0}^{\infty} \int_{[k, k+1)} 2x(1 - F(x)) dx \\
 &= \sum_{k=0}^{\infty} P(X > k) \int_{[k, k+1)} 2x dx \\
 &= \sum_{k=0}^{\infty} P(X > k)(2k + 1)
 \end{aligned}$$

since

$$\int_k^{k+1} 2x dx = x^2 \Big|_k^{k+1} = 2k + 1.$$

3. PfS Exercise 5.1.4, page 91. (a) Let $\mathcal{X} = [0, 1]$, $\mathcal{Y} = (1, \infty)$ both equipped with the Borel sets and Lebesgue measure. Let $f(x, y) = e^{-xy} - 2e^{-2xy}$. Show that
- (i) $\int_0^1 (\int_1^{\infty} f(x, y) dy) dx = \int_0^1 x^{-1} (e^{-x} - e^{-2x}) dx$ exists and is > 0 .
 - (ii) $\int_1^{\infty} (\int_0^1 f(x, y) dx) dy = \int_1^{\infty} y^{-1} (e^{-2y} - e^{-y}) dy$ exists and is < 0 .

Solution: First,

$$\int_0^1 \left(\int_1^\infty f(x, y) dy \right) dx = \int_0^1 x^{-1} (e^{-x} - e^{-2x}) dx$$

by an easy calculation of the inner integral. Note that the function $x^{-1}(e^{-x} - e^{-2x})$ converges to 1 as $x \downarrow 0$, and is continuous elsewhere, hence is uniformly continuous and uniformly bounded on $[0, 1]$. Since it is strictly positive and bounded, the last integral exists and is positive. On the other hand,

$$\int_1^\infty \left(\int_0^1 f(x, y) dx \right) dy = \int_1^\infty y^{-1} (e^{-2y} - e^{-y}) dy$$

again by an easy calculation of the inner integral. Now the integrand is negative (since $e^{-2y} < e^{-y}$ for all $y > 0$), bounded and the two integrals $\int_1^\infty y^{-1} e^{-2y} dy$ and $\int_1^\infty y^{-1} e^{-y} dy$ both are clearly finite. Thus the second iterated integral exists and is strictly negative. Letting $Ei(z) \equiv -\int_z^\infty t^{-1} e^{-t} dt$, the exponential integral function, it is easily shown that

$$\begin{aligned} \int_0^1 \left(\int_1^\infty f(x, y) dy \right) dx &= \int_0^1 x^{-1} (e^{-x} - e^{-2x}) dx \\ &= -Ei(-2) + Ei(-1) + \log(2) \\ &\doteq 0.522664\dots, \end{aligned}$$

while

$$\begin{aligned} \int_1^\infty \left(\int_0^1 f(x, y) dx \right) dy &= \int_1^\infty y^{-1} (e^{-2y} - e^{-y}) dy \\ &= -Ei(-2) + Ei(-1) \\ &\doteq -0.170483\dots \end{aligned}$$

Of course the difficulty here is that

$$\int_1^\infty \left(\int_0^1 |f(x, y)| dx \right) dy = \int_0^1 \left(\int_1^\infty |f(x, y)| dy \right) dx = \infty.$$

In fact

$$\begin{aligned} \int_0^1 \left(\int_1^\infty f^+(x, y) dx \right) dy &= \int_0^1 \left(\int_{x^{-1} \log 2}^\infty f(x, y) dy \right) dx \\ &= \int_0^1 x^{-1} (e^{-\log 2} - e^{-2 \log 2}) dx \\ &= \frac{1}{4} \int_0^1 x^{-1} dx = +\infty, \end{aligned}$$

and similarly

$$\begin{aligned} \int_0^1 \left(\int_1^\infty f^-(x, y) dx \right) dy &= \int_0^1 \left(\int_0^{x^{-1} \log 2} f(x, y) dy \right) dx \\ &= \int_0^1 x^{-1} ((1 - e^{-2 \log 2}) - (1 - e^{-\log 2})) dx \\ &= \frac{1}{4} \int_0^1 x^{-1} dx = +\infty. \end{aligned}$$

4. Show that $f(x, y) = e^{-xy} \sin(x)$ is integrable with respect on Lebesgue measure on R^2 in the strip $0 < x < a$, $0 < y$. Perform the double integral in the two different orders to find that

$$\int_0^a \frac{\sin(x)}{x} dx = \frac{\pi}{2} - \cos(a) \int_0^\infty \frac{e^{-ay}}{1+y^2} dy - \sin(a) \int_0^\infty \frac{ye^{-ay}}{1+y^2} dy.$$

Use the inequality $1 + y^2 \geq 1$ to obtain the bound

$$\left| \int_0^a \frac{\sin(x)}{x} dx - \frac{\pi}{2} \right| \leq \frac{2}{a}.$$

Letting $a \rightarrow \infty$ yields $\int_0^\infty x^{-1} \sin(x) dx = \pi/2$.

Solution: Note that

$$\int_0^a \int_0^\infty e^{-xy} |\sin(x)| dy dx = \int_0^a |\sin(x)| x^{-1} dx \leq a < \infty$$

since $|\sin(x)| \leq x$, $x > 0$. Hence we may apply Fubini's theorem to compute the double integral in two different ways: the first iterated integral, integrating on y first yields

$$\int_0^a \left(\int_0^\infty \sin(x)e^{-xy} dy \right) dx = \int_0^a \sin(x) \frac{1}{x} dx.$$

The second iterated integral, integrating on x first yields

$$\begin{aligned} & \int_0^\infty \left(\int_0^a \sin(x)e^{-xy} dx \right) dy \\ &= \int_0^\infty \left(\frac{1}{1+y^2} - \frac{e^{-ay}(\cos(a) + y \sin(a))}{1+y^2} \right) dy \\ &= \frac{\pi}{2} - \cos(a) \int_0^\infty \frac{e^{-ay}}{1+y^2} dy - \sin(a) \int_0^\infty \frac{ye^{-ay}}{1+y^2} dy. \end{aligned}$$

Because these two iterated integrals are equal by Fubini's theorem, it follows, using $1+y^2 \geq 1$ for $y \geq 0$, that, for $a \geq 1$,

$$\left| \int_0^a \frac{\sin(x)}{x} dx - \frac{\pi}{2} \right| \leq \int_0^\infty e^{-ay} dy + \int_0^\infty ye^{-ay} dy \leq \frac{2}{a}.$$

Letting $a \rightarrow \infty$ yields $\int_0^\infty x^{-1} \sin(x) dx = \pi/2$.

5. **Optional bonus problem:** Suppose that $\phi : (0, 1) \rightarrow R^+$ satisfies:

- (i) $\phi(1) = 0$,
- (ii) $\phi' < 0$,
- (iii) $\phi'' > 0$.
- (iv) $\phi(u) \rightarrow \infty$ as $u \downarrow 0$.

For example: (a) $\phi(u) = (u^{-\theta} - 1)/\theta$, $0 < \theta < \infty$; (b) $\phi(u) = (-\log u)^{1/\theta}$, $0 < \theta < 1$; (c) $\phi(u) = \log(\frac{1-\theta}{1-\theta u})$, $0 < \theta < \infty$, $\theta \neq 1$.

Define $C : (0, 1] \times (0, 1] \rightarrow [0, 1]$ by $C(u, v) = \phi^{-1}(\phi(u) + \phi(v))$.

A. Show that C is the joint distribution function of a pair of random variables (U, V) which have the property that the marginal distributions of U and V are Uniform(0, 1).

B. Show that under (i) - (iv) (U, V) have a joint density function $c(u, v)$ and calculate it; i.e. show that $C(u, v) \ll \text{Lebesgue measure on } [0, 1]^2$.

C. What if $\lim_{u \rightarrow 0} \phi(u) = \phi(0) < \infty$? Does (U, V) have a joint density function? Show that if $\lim_{u \rightarrow 0} \phi(u) = \phi(0) < \infty$, then C is still a distribution function, but with one component C_{ac} with a density function

c_{ac} (“ac” for “absolutely continuous with respect to Lebesgue measure λ_2 on R^2 ”), and a component which is singular with respect to λ_2 . Show that the singular component has probability $p = -\frac{\phi(0)}{\phi'(0)}$ on the curve given by: $\phi(u) + \phi(v) = \phi(0)$.

Solution: A. Note that $C(u, v) = \phi^{-1}(\phi(u) + \phi(v))$ for (u, v) such that $\phi(u) + \phi(v) \leq \phi(0)$, and is naturally defined by $C(u, v) = 0$ for (u, v) such that $\phi(u) + \phi(v) > \phi(0)$. For (u, v) such that $C(u, v) > 0$, we can calculate the joint density as follows: since $\phi(C(u, v)) = \phi(u) + \phi(v)$ on $D \equiv \{(u, v) \in (0, 1)^2 : C(u, v) > 0\}$, it follows that

$$\phi'(C(u, v)) \frac{\partial}{\partial u} C(u, v) = \phi'(u)$$

and

$$\phi'(C(u, v)) \frac{\partial}{\partial v} C(u, v) = \phi'(v).$$

Then we also have

$$\phi''(C(u, v)) \frac{\partial}{\partial u} C(u, v) \frac{\partial}{\partial v} C(u, v) + \phi'(C(u, v)) \frac{\partial^2}{\partial u \partial v} C(u, v) = 0.$$

Solving the last equation for $\frac{\partial^2}{\partial u \partial v} C(u, v) \equiv c(u, v)$ yields, upon use of the preceding two displays,

$$c(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v) = -\frac{\phi''(C(u, v))\phi'(u)\phi'(v)}{[\phi'(C(u, v))]^3}.$$

From the properties of ϕ it follows that $c(u, v)$ is positive on the set D , and hence for any rectangle $A \subset D$ we have $C(A) = \int \int_D c(u, v) du dv \geq 0$. Also note that $C(u, 1) = u$ for $u \in (0, 1)$, and $C(1, v) = v$ for $v \in (0, 1)$, and furthermore $C(1, 1) = 1$. Also C is continuous from above by virtue of the continuity of the function ϕ . Hence C is a bivariate distribution function on $[0, 1]^2$ with uniform marginal distributions.

B. When $\phi(u) \rightarrow \infty$ as $u \downarrow 0$, $C(u, v)$ is positive for all $(u, v) \in (0, 1)^2$, and the density $c(u, v)$ calculated in part A works for all $(u, v) \in (0, 1)^2$.

C. When $\phi(0) < \infty$, then the mass p accounted for by the density c is

$$\begin{aligned}
p &= \iint_{\phi(u)+\phi(v)<\phi(0)} c(u,v)dudv \\
&= - \iint_{\phi(u)+\phi(v)<\phi(0)} \frac{\phi''(C(u,v))\phi'(u)\phi'(v)}{[\phi'(C(u,v))]^3} dudv \\
&= - \int \int_{0 < u < v < 1} \frac{\phi''(u)}{[\phi'(u)]^2} \phi'(v) dv \\
&= \int_0^1 \frac{\phi''(u)}{[\phi'(u)]^2} \phi(u) du \\
&= - \left[\frac{\phi(u)}{\phi'(u)} \right]_0^1 + 1 = \frac{\phi(0)}{\phi'(0)} + 1
\end{aligned}$$

where the next to last equality follows from integration by parts, and the last equality follows from $\lim_{a \rightarrow 1} \phi(a)/\phi'(a) = 0$ by convexity of ϕ since $\phi(a)/\phi'(a)$ is the x intercept of the tangent line to ϕ at $a < 1$ (draw the picture!). Thus $p < 1$ if and only if $\phi(0)/\phi'(0) \neq 0$. In the latter case the distribution C has a singular component on the boundary curve $B \equiv \{(u, v) \in (0, 1)^2 : \phi(u) + \phi(v) = \phi(0)\}$, and the probability that $(U, V) \in B$ is exactly $-\phi(0)/\phi'(0)$.

Questions: what is the length of the boundary curve? Is there a “density” for the remaining mass with respect to the natural distance measure along this curve?

Note 1: This problem is based on

- Genest, C. and MacKay, J. (1986). The joy of copulas: bivariate distributions with uniform marginals. *The American Statistician* **40**, 280 - 283.

Note 2: Further results on allowable functions ϕ for multivariate Archimedean copula is given by

- McNeil, A. J., Neslehová, J. (2009). Multivariate Archimedean copulas, d-monotone functions and l_1 -norm symmetric distributions. *Ann. Statist.* **37**, 3059 - 3097.