Probability & Statistics for EECS: Homework #7 Solution

Problem 1

Let

$$F(x) = \frac{2}{\pi} \sin^{-1}(\sqrt{x}), \text{ for } 0 < x < 1,$$

and let $F(x) = 0$ for $x \le 0$ and $F(x) = 1$ for $x \ge 1$.

(a) Check that F is a valid CDF, and find the corresponding PDF f.

(b) Explain how it is possible for f to be a valid PDF even though $f(x)$ goes to ∞ as x approaches 0 and as x approaches 1.

Solution

(a) The function F is increasing since the square root and $\sin^{-1}$ functions are increasing. It is continuous since $\frac{2}{\pi}\sin^{-1}(\sqrt{0}) = 0$, $\frac{2}{\pi}\sin^{-1}(\sqrt{1}) = 1$, the square root function is continuous on $(0, \infty)$, the $\sin^{-1}$ function is continuous on $(-1, 1)$, and a constant function is continuous everywhere. And $F(x) \to 0$ as $x \to -\infty$, $F(x) \to 1$ as $x \to \infty$ (since in fact $F(x) = 0$ for $x \le 0$ and $F(x) = 1$ for $x \ge 1$). So F is a valid CDF. By the chain rule, the corresponding PDF is

$$f(x) = \frac{2}{\pi} \cdot \frac{1}{2\sqrt{x}} \cdot \frac{1}{\sqrt{1-x}} = \frac{1}{\pi} x^{-1/2}(1-x)^{-1/2}$$

for $0 < x < 1$ (and 0 otherwise).

(b) By (a), $f(x) \to \infty$ as $x \to 0$, and also $f(x) \to \infty$ as $x \to 1$. But the area under the curve is still finite (in particular, the area is 1). There is no contradiction in this. For a simpler example, note that

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{a \to 0^+} \int_a^1 \frac{1}{\sqrt{x}} dx = \lim_{a \to 0^+} (2\sqrt{1} - 2\sqrt{a}) = 2$$

which is finite even though $1/\sqrt{x} \to \infty$ as x approaches 0 from the right.

Problem 2

Let F be a CDF which is continuous and strictly increasing. Let μ be the mean of the distribution. The quantile function, F^{-1}, has many applications in statistics and econometrics. Show that the area under the curve of the quantile function from 0 to 1 is μ.

Solution

Solution: We want to find $\int_0^1 F^{-1}(u)du$. Let $U \sim \text{Unif}(0,1)$ and $X = F^{-1}(U)$. By universality of the Uniform, $X \sim F$. By LOTUS,

$$\int_0^1 F^{-1}(u)du = E\left(F^{-1}(U)\right) = E(X) = \mu$$

Equivalently, make the substitution $u = F(x)$, so $du = f(x)dx$, where f is the PDF of the distribution with CDF F. Then the integral becomes

$$\int_{-\infty}^{\infty} F^{-1}(F(x))f(x)dx = \int_{-\infty}^{\infty} xf(x)dx = \mu.$$

Sanity check: For the simple case that F is the $\text{Unif}(0,1)$ CDF, which is $F(u) = u$ on $(0,1)$, we have $\int_0^1 F^{-1}(u)du = \int_0^1 udu = 1/2$, which is the mean of a $\text{Unif}(0,1)$.

Problem 3

Let $U_1, \ldots, U_n$ be i.i.d. Unif$(0,1)$, and $X = \max(U_1, ..., U_n)$. What is the PDF of X? What is $E(X)$?

Solution

Solution: Note that $X \leq x$ holds if and only if all of the U_j's are at most x . So the CDF of X is

$$P(X \leq x) = P\left(U_1 \leq x, U_2 \leq x, \ldots, U_n \leq x\right) = \left(P\left(U_1 \leq x\right)\right)^n = x^n$$

for $0 < x < 1$ (and the CDF is 0 for $x \leq 0$ and 1 for $x \geq 1$). So the PDF of X is

$$f(x) = nx^{n-1}$$

for $0 < x < 1$ (and 0 otherwise). Then

$$E(X) = \int_0^1 x\left(nx^{n-1}\right) dx = n \int_0^1 x^n dx = \frac{n}{n+1}.$$

(For generalizations of these results, see the material on order statistics in Chapter 8.)

Problem 4

A stick of length 1 is broken at a uniformly random point, yielding two pieces. Let X and Y be the lengths of the shorter and longer pieces, respectively, and let $R = X/Y$ be the ratio of the lengths X and Y.

(a) Find the CDF and PDF of R.

(b) Find the expected value of R (if it exists).

(c) Find the expected value of $1/R$ (if it exists).

Solution

(a) Let $U \sim \text{Unif}(0,1)$ be the break point, so $X = \min(U, 1-U)$. For $r \in (0,1)$,

$$P(R \le r) = P(X \le r(1-X)) = P\left(X \le \frac{r}{1+r}\right)$$

This is the CDF of X, evaluated at $r/(1+r)$, so we just need to find the CDF of X:

$$P(X \le x) = 1 - P(X > x) = 1 - P(U > x, 1-U > x) = 1 - P(x < U < 1-x) = 2x$$

for $0 \le x \le 1/2$, and the CDF is 0 for $x < 0$ and 1 for $x > 1/2$. So $X \sim \text{Unif}(0, 1/2)$, and the CDF of R is

$$P(R \le r) = P\left(X \le \frac{r}{1+r}\right) = \frac{2r}{1+r}$$

for $r \in (0,1)$, and it is 0 for $r \le 0$ and 1 for $r \ge 1$. Alternatively, we can note that

$$P\left(X \le \frac{r}{1+r}\right) = P\left(U \le \frac{r}{1+r} \text{ or } 1-U \le \frac{r}{1+r}\right).$$

The events $U \le r/(1+r)$ and $1-U \le r/(1+r)$ are disjoint for $r \in (0,1)$ since the latter is equivalent to $U \ge 1/(1+r)$. Thus, again for $r \in (0,1)$ we have

$$P(R \le r) = P\left(U \le \frac{r}{1+r}\right) + P\left(1-U \le \frac{r}{1+r}\right) = \frac{2r}{1+r}.$$

The PDF is 0 if r is not in $(0,1)$, and for $r \in (0,1)$ it is

$$f(r) = \frac{2(1+r) - 2r}{(1+r)^2} = \frac{2}{(1+r)^2}$$

(b) We have

$$E(R) = 2\int_0^1 \frac{r}{(1+r)^2} dr = 2\int_1^2 \frac{(t-1)}{t^2} dt = 2\int_1^2 \frac{1}{t} dt - 2\int_1^2 \frac{1}{t^2} dt = 2\ln 2 - 1.$$

(c) This expected value does not exist, since $\int_0^1 \frac{1}{r(1+r)^2} dr$ diverges. To show this, note that $\int_0^1 \frac{1}{r} dr$ diverges and $\frac{1}{r(1+r)^2} \ge \frac{1}{4r}$ for $0 < r < 1$.

Problem 5

The Exponential is the analog of the Geometric in continuous time. This problem explores the connection between Exponential and Geometric in more detail, asking what happens to a Geometric in a limit where the Bernoulli trials are performed faster and faster but with smaller and smaller success probabilities.
Suppose that Bernoulli trials are being performed in continuous time; rather than only thinking about first trial, second trial, etc., imagine that the trials take place at points on a timeline. Assume that the trials are at regularly spaced times 0, Δt, $2\Delta t$,... , where Δt is a small positive number. Let the probability of success of each trial be $\lambda\Delta t$, where λ is a positive constant. Let G be the number of failures before the first success (in discrete time), and T be the time of the first success (in continuous time).

(a) Find a simple equation relating G to T. Hint: Draw a timeline and try out a simple example.

(b) Find the CDF of T. Hint: First find $P(T > t)$.

(c) Show that as $\Delta t \to 0$, the CDF of T converges to the Expo(λ) CDF, evaluating all the CDFs at a fixed $t \geq 0$.

Solution

(a) $T = G\Delta t$.

(b) For $t \geq 0$, $P\left(T > t\right) = P\left(G > \frac{t}{\Delta t}\right) = P\left(\text{no success in the first } \left\lfloor \frac{t}{\Delta t} \right\rfloor \text{ trials}\right) = \left(1 - \lambda\Delta t\right)^{\left\lfloor \frac{t}{\Delta t} \right\rfloor}$. Thus The CDF of T is

$$P\left(T \leq t\right) = 1 - P\left(T > t\right) = 1 - \left(1 - \lambda\Delta t\right)^{\left\lfloor \frac{t}{\Delta t} \right\rfloor}.$$

(c) As $\Delta t \to 0$,

$$\begin{aligned} \lim_{\Delta t \to 0} P\left(T \leq t\right) &= \lim_{\Delta t \to 0} \left[1 - \left(1 - \lambda\Delta t\right)^{\left\lfloor \frac{t}{\Delta t} \right\rfloor}\right] = 1 - \lim_{\Delta t \to 0} \left(1 - \lambda\Delta t\right)^{\frac{t}{\Delta t}} \\ &= 1 - \lim_{\Delta t \to 0} \left[\left(1 - \lambda\Delta t\right)^{\frac{1}{\lambda\Delta t}}\right]^{\lambda t} = 1 - e^{-\lambda t}. \end{aligned}$$

Thus for $t \geq 0$, the CDF of T converges to the Expo(λ) CDF as $\Delta t \to 0$.

Problem 6

Let $Z \sim \mathcal{N}(0,1)$, and c be a nonnegative constant. Find $E(\max(Z-c,0))$, in terms of the standard Normal CDF Φ and PDF φ.

Solution

Hint: Use LOTUS, and handle the max symbol by adjusting the limits of integration appropriately. As a check, make sure that your answer reduces to $1/\sqrt{2\pi}$ when $c=0$; this must be the case since we show in Chapter 7 that $E|Z| = \sqrt{2/\pi}$, and we have $|Z| = \max(Z,0) + \max(-Z,0)$ so by symmetry

$$E|Z| = E(\max(Z,0)) + E(\max(-Z,0)) = 2E(\max(Z,0)).$$

Solution: Let φ be the $\mathcal{N}(0,1)$ PDF. By LOTUS,

$$\begin{aligned}
E(\max(Z-c,0)) &= \int_{-\infty}^{\infty} \max(z-c,0)\varphi(z)dz \\
&= \int_{c}^{\infty} (z-c)\varphi(z)dz \\
&= \int_{c}^{\infty} z\varphi(z)dz - c\int_{c}^{\infty} \varphi(z)dz \\
&= \frac{-1}{\sqrt{2\pi}} e^{-z^2/2}\Big|_{c}^{\infty} - c(1-\Phi(c)) \\
&= \frac{1}{\sqrt{2\pi}} e^{-c^2/2} - c(1-\Phi(c)).
\end{aligned}$$