Chapter 2: Random Variables

All of Statistics, Wasserman

2. Let X be such that $P(X=2)=P(X=3)=1/10$ and $P(X=5)=8/10$. Find the CDF and evaluate $P(2<X\leq 4.8)$ and $P(2\leq X\leq 4.8)$.

Solution. By definition,

$$F(x)=P(X\leq x)=\begin{cases}0 & \text{if } x<2,\\ 1/10 & \text{if } 2\leq x<3,\\ 2/10 & \text{if } 3\leq x<5,\\ 1 & \text{if } x\geq 5.\end{cases}$$

The only x with $2<x\leq 4.8$ and non-zero probability is $x=3$. Then $P(2<X\leq 4.8)=P(X=3)=1/10$. Analogously, $P(2\leq X\leq 4.8)=P(X=2)+P(X=3)=2/10$. $\diamondsuit$

3. Let F be the CDF for a random variable X. Then:

(a) $P(X=x)=F(x)-F(x^-)$ where $F(x^-)=\lim_{y\uparrow x}F(y)$;

(b) $P(x<X\leq y)=F(y)-F(x)$;

(c) $P(X>x)=1-F(x)$;

(d) If X is continuous, then

$$F(b)-F(a)=P(a<X<b)=P(a\leq X<b)=P(a<X\leq b)=P(a\leq X\leq b).$$

Proof. (a) Clearly, $\lim_{y\uparrow x}\{X\leq y\}=\{X<x\}$. Then (by the continuity of probabilities theorem)

$$F(x^-)=\lim_{y\uparrow x}P(X\leq y)=P(X<x)=P(X\leq x)-P(X=x)=F(x)-P(X=x).$$

(b) Since $\{X\leq x\}\sqcup\{x<X\leq y\}=\{X\leq y\}$, then

$$P(x<X\leq y)=P(X\leq y)-P(X\leq x)=F(y)-F(x).$$

(c) Since $\Omega=\{X>x\}\sqcup\{X\leq x\}$, then

$$1=P(\Omega)=P(X>x)+P(X\leq x)=P(X>x)+F(x).$$

(d) If X is continuous, $P(X=x)=0$ for all x. Using this and expressing the intervals as disjoint unions (such as $\{a\leq X<b\}=\{X=a\}\sqcup\{a<X<b\}$) we obtain the desired result.

$\square$

4. Let X have pdf

$$f(x)=\begin{cases}1/4 & \text{if } 0<x<1,\\ 3/8 & \text{if } 3<x<5,\\ 0 & \text{otherwise.}\end{cases}$$

(a) Find the CDF of X.

(b) Let $Y = 1/X$. Find the pdf of Y.

Solution. (a) By definition,

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & \text{if } x \le 0, \\ (1/4)x & \text{if } 0 < x < 1, \\ 1/4 & \text{if } 1 \le x \le 3, \\ 1/4 + (3/8)(x-3) & \text{if } 3 < x < 5, \\ 1 & \text{if } x \ge 5. \end{cases}$$

(b) Since $r = 1/x$ is monotone decreasing in $(0, \infty)$ it has an inverse $s = 1/y$ and

$$f_Y(y) = f_X(s(y)) \left| \frac{ds(y)}{dy} \right| = \begin{cases} \frac{3}{8y^2} & \text{if } 1/5 < y < 1/3, \\ \frac{1}{4y^2} & \text{if } y > 1, \\ 0 & \text{otherwise.} \end{cases}$$

◊

5. Let X and Y be discrete random variables. Show that X and Y are independent if and only if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \tag{1}$$

for all x, y.

Proof. If X and Y are independent, eq. (1) is trivial. Conversely, suppose eq. (1). Then for any A, B,

$$P(X \in A,\, Y \in B) = \sum_{x \in A} \sum_{y \in B} f(x, y) = \sum_{x \in A} \sum_{y \in B} f(x) f(y) =$$

$$\sum_{x \in A} \left(f(x) \sum_{y \in B} f(y) \right) = \left(\sum_{x \in A} f(x) \right) \left(\sum_{y \in B} f(y) \right) = P(X \in A) P(Y \in B).$$

□

6. Let X have distribution F and density f. Let $A \subset \mathbb{R}$ and let $Y = \chi_A(X)$ where χ_A is the characteristic function of A. Find the CDF of Y.

Solution. By definition, Y has a Bernoulli distribution with success probability

$$P(Y = 1) = P(\chi_A(X) = 1) = P(X \in A) = \int_A f(x)\,dx.$$

Therefore,

$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0, \\ 1 - \int_A f(x)\,dx & \text{if } 0 \le y < 1, \\ 1 & \text{if } y \ge 1. \end{cases}$$

◊

7. Let X and Y be independent and suppose that each has Uniform$(0, 1)$ distribution. Let $Z = \min\{X, Y\}$. Find the density of Z.

Solution. Following the hint,

$$P(Z > z) = P(\min\{X, Y\} > z) = P(X > z,\, Y > z) = P(X > z)P(Y > z) = (1 - F_X(z))(1 - F_Y(z))$$

where the third equality is because of independence. Furthemore,

$$P(Z \le z) = 1 - (1 - F_X(z))(1 - F_Y(z))$$

Then, since the CDF of a Uniform$(0, 1)$ distributed random variable is the identity in $[0, 1]$,

$$F_Z(z) = 1 - (1 - z)^2 = 2z - z^2 \text{ for } 0 \le x \le 1$$

In particular,

$$f_Z(z) = 2 - 2z \text{ for } 0 \le x \le 1.$$

$\Diamond$

8. Let X have CDF F. Find the CDF of $X^+ = \max\{0, X\}$.

Solution. We distinguish two cases:

$$P(X^+ \le z) = \begin{cases} P(X \le z) & \text{if } z \ge 0, \\ 0 & \text{otherwise.} \end{cases}$$

$\Diamond$

10. Let X and Y be independent. Prove that $g(X)$ and $h(Y)$ are independent, where g and h are functions.

Proof. Simply use $\{g(X) \in A\} = \{X \in g^{-1}(A)\}$. $\square$

11. (a) Suppose we toss a coin once and let p be the probability of heads. Let X and Y be the number of heads and tails. Prove that X and Y are dependent.

(b) Let $N \sim \text{Poisson}(\lambda)$ and suppose we toss a coin N times. Let X and Y be the number of heads and tails. Show that X and Y are independent.

Proof. (a) Note that

$$P(X = 1,\, Y = 1) = 0 < P(X = 1)P(Y = 1).$$

(b) Calculating directly,

$$P(X = x,\, Y = y) = P(X = x \mid N = x + y)P(N = x + y) = \left(\binom{x+y}{x} p^x (1-p)^y\right)\left(e^{-\lambda}\frac{\lambda^{x+y}}{(x+y)!}\right)$$

since $X \mid N \sim \text{Binomial}(N, p)$. This expression may be factored as a product of the form $g(x)h(y)$ (the $(x + y)!$'s will cancel each other) and therefore, X and Y are independent.

$\square$

12. Suppose the range of X and Y is a (possibly infinite) rectangle. If $f_{X,Y}(x, y) = g(x)h(y)$ for some functions g and h (not necessarily pdfs) then X and Y are independent.

Proof. Note that

$$f_X(x) = \int f_{X,Y}(x, y)\, dy = \int g(x)h(y)\, dy = g(x) \int h(y)\, dy.$$

Analogously,

$$f_Y(y) = h(y) \int g(x)\, dx.$$

Therefore,

$$f_X(x)f_Y(y) = \left(g(x)\int h(y)\,dy\right)\left(h(y)\int g(x)\,dx\right) = g(x)h(y)$$

where the last equality follows from $1 = \iint f(x,y)\,dxdy = \int g(x)\,dx \int h(y)\,dy$ (which is possible because the domain is a rectangle and because of Fubini). $\square$

13. Let $X \sim N(0,1)$ and let $Y = e^X$. Find the pdf of Y.

Proof. Since $r = e^x$ is strictly increasing in $\mathbb{R}$, it has an inverse $s = \log y$ and

$$f_Y(y) = f_X(s(y))\left|\frac{ds(y)}{dy}\right| = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(\log y)^2}{2}\right)\left|\frac{1}{y}\right|.$$

$\square$

14. Let (X,Y) be uniformly distributed on the unit disk $D^2 = \{(x,y) : x^2+y^2 \leq 1\}$. Let $R = \sqrt{X^2+Y^2}$. Find the CDF and PDF of R.

Proof. Note that

$$A_r = \{(x,y) \in D^2 : \sqrt{x^2+y^2} \leq r\} = \begin{cases} \emptyset & \text{if } r \leq 0, \\ B_r(0) & \text{if } 0 < r < 1, \\ D^2 & \text{if } r \geq 1 \end{cases}$$

Therefore,

$$\begin{aligned} F_R(r) &= \iint_{A_r} f(x,y)\,dxdy \\ &= \begin{cases} \iint_\emptyset 1/\pi\,dxdy & \text{if } r \leq 0, \\ \iint_{B_r(0)} 1/\pi\,dxdy & \text{if } 0 < r < 1, \\ \iint_{D^2} 1/\pi\,dxdy & \text{if } r \geq 1. \end{cases} \\ &= \begin{cases} 0 & \text{if } r \leq 0, \\ r^2 & \text{if } 0 < r < 1, \\ 1 & \text{if } r \geq 1. \end{cases} \end{aligned}$$

Furthermore,

$$f_R(r) = 2r \text{ for } 0 < r < 1.$$

$\square$

15. Let X have a coninuous, strictly increasing CDF F. Let $Y = F(X)$. Find the density of Y. Now let $U \sim \text{Uniform}(0,1)$ and let $X = F^{-1}(U)$. Show that $X \sim F$. (This result still holds without the continuous, strictly increasing hypothesis.)

Proof. Note that Y only takes values in $[0,1]$. Then $F_Y(y) = 0$ if $y \leq 0$ and $F_Y(y) = 1$ if $y \geq 1$ and

$$F_Y(y) = P(Y \leq y) = P(F(X) \leq y) = P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y \text{ if } 0 < y < 1.$$

Therefore,

$$f(y) = 1 \text{ for } 0 < y < 1$$

and $Y \sim \text{Uniform}(0,1)$. Conversely,

$$F_X(x) = P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x).$$

That is, $X \sim F$. $\square$

16. Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ and assume X and Y are independent. Show that the distribution of X given $X+Y=n$ is $\text{Binomial}(n,\pi)$ where $\pi = \lambda/(\lambda+\mu)$.

Proof. Following the hint, for $x \in \mathbb{N}$,

$$\begin{aligned}
P(X = x \mid X+Y=n) &= \frac{P(X=x,\, X+Y=n)}{P(X+Y=n)} \\
&= \frac{P(X=x,\, Y=n-x)}{P(X+Y=n)} \\
&= \frac{\left(e^{-\lambda}\frac{\lambda^x}{x!}\right)\left(e^{-\mu}\frac{\mu^{n-x}}{(n-x)!}\right)}{e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^n}{n!}} \\
&= \frac{\left(\frac{\lambda^x}{x!}\right)\left(\frac{\mu^{n-x}}{(n-x)!}\right)}{\frac{(\lambda+\mu)^n}{n!}} \\
&= \binom{n}{x}\pi^x(1-\pi)^{n-x}.
\end{aligned}$$

Remark. The trick $\{X=x,\, X+Y=n\} = \{X=x,\, Y=n-x\}$ looks very useful. ☐

17. Let

$$f_{X,Y}(x,y) = \begin{cases} c(x+y^2) & \text{if } 0 \le x \le 1 \text{ and } 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find $P(X < 1/2 \mid Y = 1/2)$.

Solution. By definition,

$$f_Y(y) = \int f_{X,Y}(x,y)\,dx = c\int_0^1 (x+y^2)\,dx = c(1/2+y^2).$$

Then

$$f(x \mid y) = \frac{f(x,y)}{f(y)} = \frac{x+y^2}{1/2+y^2}.$$

Finally,

$$P(X<1/2 \mid Y=1/2) = \int_0^{1/2} f(x \mid 1/2)\,dx = \int_0^{1/2} \frac{x+1/4}{1/2+1/4}\,dx = 1/3.$$

◊

19. Let X be continuous and $Y = r(X)$. If r is strictly increasing or decreasing, then r has an inverse $s = r^{-1}$ and

$$f_Y(y) = f_X(s(y))\left|\frac{ds(y)}{dy}\right|.$$

Proof. Suppose r is strictly decreasing. Then

$$F_Y(y) = P(Y \le y) = P(r(X) \le y) = \\ P(X \ge r^{-1}(y)) = 1 - P(X < r^{-1}(y)) = 1 - F_X(r^{-1}(y)).$$

Differentiating yields,

$$f_Y(y) = -f_X(s(y))\frac{ds(y)}{dy} = f_X(s(y))\left|\frac{ds(y)}{dy}\right|$$

where the last equality is because s is decreasing (since r is too). The other case is analogous. ☐