STA 200A: Homework 6

Note: Below the notation 3.T11 means Chapter 3, Theoretical Exercise 11. Similarly, the notation 4.P21 means Chapter 4, Problem 21.

1. 5.T29

Solution: Let us assume that X has positive pdf over a region (a, b) including $a = -\infty, b = \infty$. Then F is strictly increasing and the pdf of Y is $f(F^{-1}(Y))(F^{-1})'(Y) = f(F^{-1}(Y))/f(F^{-1}(Y)) = 1, 0 \leq Y \leq 1$. We could also see it by $F_Y(y) = P\{Y \leq y\} = P\{F(X) \leq y\} = P\{X \leq F^{-1}(y)\} = F(F^{-1}(y)) = y$ for $0 \leq y \leq 1$.

2. 6.P10

Solution:

(a) By symmetry $P\{X < Y\} = P\{Y < X\}$ and $P\{Y < X\} + P\{X < Y\} = 1$. Hence, $P\{X < Y\} = 1/2$.

(b) Then

$$P\{X < a\} = \int_0^a \int_0^\infty e^{-(x+y)} dy dx = \int_0^a e^{-x} dx = 1 - e^{-a}.$$

3. 6.P14

Solution: Let X denote the random variable providing the location of the accident along the road. Then according to the problem specification, let X be uniformly distributed between $[0, L]$. Let Y denote the random variable the location of the ambulance. Then we define $D = |X - Y|$ the random variable representing the distance between the accident and the ambulance. We want to compute

$$P(D \leq d) = \int \int_{X,Y \in \Omega} f(x, y) dx dy,$$

with Ω the set of points where $|X - Y| \leq d$.

The above integral is then

$$\begin{aligned}
P(D \leq d) &= \int_{x=0}^{d} \int_{y=0}^{x+d} f(x, y) dy dx + \int_{x=d}^{L-d} \int_{y=x-d}^{x+d} f(x, y) dy dx + \int_{x=L-d}^{L} \int_{y=x-d}^{L} f(x, y) dy dx \\
&= \frac{1}{L^2} \int_0^d (x + d) dx + \frac{1}{L^2} \int_d^{L-d} (x + d - (x - d)) dx + \frac{1}{L^2} \int_{L-d}^{L} (L - x + d) dx \\
&= \frac{(2L - d)d}{L^2}
\end{aligned}$$

Then take derivative on both sides with respect to d, we have $f_D(d) = \frac{2(L-d)}{L^2}$.

4. 6.P15

Solution:

(a) Let $|R|$ be the area of R.

$$\int_R f(x,y)dxdy = c|R| = 1$$

(b) Let $1(A)$ be 1 if A is true and 0 otherwise. The area is $2^2 = 4$ so density is

$$f(x,y) = \frac{1}{4}1\{|x| < 1, |y| < 1\}.$$

and hence,

$$f(x) = \frac{1}{2}1\{|x| < 1\}, \quad f(y) = \frac{1}{2}1\{|y| < 1\}$$

which implies

$$f(x,y) = f(x)f(y),$$

and they are indeed independent.

(c) The probability of $x^2 + y^2 < 1$ is the area of the unit circle divided by 4, $\pi/4$.

5. 6.P20

Solution: The density can be factorized by

$$f(x,y) = \left(xe^{-x}1\{x > 0\}\right)\left(e^{-y}1\{y > 0\}\right)$$

which means that they are independent. Also, if

$$f(x,y) = 21\{0 < x < y, 0 < y < 1\}$$

then we have that $Y < 1/2$ implies that $X < 1/2$ with probability 1 while marginally $X > 1/2$ has a non-zero probability.

6. 6.P22

Solution:

(a) The density does not factorize so they are not independent.

(b)

$$\int_0^1 (x+y)dy = x + \frac{1}{2}$$

for $0 \le x \le 1$.

(c)

$$P\{X+Y < 1\} = \int_0^1 \int_0^{1-x} (x+y)dydx = \int_0^1 (x(1-x) + \frac{(1-x)^2}{2})dx = \int_0^1 (\frac{1}{2}(-x^2+1))dx = \frac{1}{3}$$

7. Monthly sales are independent normal random variables with mean 100 and standard deviation 5.
(a) Find the probability that exactly 3 of the next 6 months have sales greater than 100.
(b) Find the probability that the total of the sales in the next 4 months is greater than 420.

Solution:

(a) The probability of a month sales being greater than 100 is 1/2. Then the probability that the number of months that exceed 100 is 3 is $\binom{6}{3}/2^6 = 0.3125$.

(b) The total sales has mean 400 and variance $4(25) = 100$ and standard deviation 10. Then the Z-score of 420 is 2, which has probability of exceedence, .0228.

8. 6.T22

Solution:

$$\begin{aligned} f_{W|X_i,i=1,\ldots,n}(w|x_1,\ldots,x_n) &= \frac{f(x_1,\ldots,x_n|w)f_w(w)}{f(x_1,\ldots,x_n)} \\ &= C\prod_{i=1}^{n} we^{-wx_i}e^{-\beta w}(\beta w)^{t-1} \\ &= Ke^{-w\left(\beta+\sum_{i=1}^{n} x_i\right)}w^{n+t-1}, \end{aligned}$$

where C and K are two normalizing constants.

9. An insurance company supposes that each person has an "accident parameter" λ. The number of accidents that someone has each year is assumed to be a random variable with a Poisson(λ) distribution. The company also assumes that the λ value of a newly insured person can be treated as random, where λ follows a Gamma distribution with parameters s and α. If a newly insured person has n accidents in her first year, find the conditional density of her accident parameter.

Solution: This is a problem in so-called Bayesian inference. By this what I mean is that given information about the number of accidents that have occurred, we want to find the density (given this number of accidents) of the accident rate λ. Before observing the number of accidents in a year, the unknown λ is governed by a gamma distribution with parameters s and α. Specifically,

$$f(\lambda) = \frac{se^{s\lambda}(s\lambda)^{\alpha-1}}{\Gamma(\alpha)}1_{\{\lambda\geq 0\}}$$

Using the Bayes rule, we have

$$\begin{aligned} P(\lambda|N=n) &= \frac{p(N=n|\lambda)f(\lambda)}{\int_\Lambda p(N=n|\lambda)f(\lambda)d\lambda} \\ &\propto p(N=n|\lambda)f(\lambda) \\ &\propto \left(\frac{e^{-\lambda}\lambda^n}{\Gamma(n+1)}\right)\left(\frac{se^{s\lambda}(s\lambda)^{\alpha-1}}{\Gamma(\alpha)}\right) \\ &\propto e^{-(1+s)\lambda}\lambda^{n+\alpha+1}. \end{aligned}$$

Now we recognize that it is a gamma distribution with parameter $(n+\alpha, 1+s)$. So

$$p(\lambda|N=n) = \frac{(s+1)^{n+\alpha}\lambda^{n+\alpha-1}e^{-(s+1)\lambda}}{\Gamma(n+\alpha)}1_{\{\lambda\geq 0\}}.$$