STAT 510: Homework 01

David Dalpiaz

Due: Monday, January 31, 11:59 PM

Exercise 1 (Independent Events)

Let A and B be independent events. Show that the following are independent:

- A^c and B
- A and B^c
- A^c and B^c

Solution

Since A and B are independent we may use

$$P(AB) = P(A) \cdot P(B)$$

First we have,

$$\begin{aligned} P(A^c) \cdot P(B) &= (1 - P(A)) \cdot P(B) \\ &= P(B) - P(A) \cdot P(B) \\ &= P(B) - P(AB) \\ &= P(A^c B) \quad \blacksquare \end{aligned}$$

Next, for A and B^c we show

$$\begin{aligned} P(A) \cdot P(B^c) &= P(A) \cdot (1 - P(B)) \\ &= P(A) - P(A) \cdot P(B) \\ &= P(A) - P(AB) \\ &= P(AB^c) \quad \blacksquare \end{aligned}$$

Lastly, we finish with

$$\begin{aligned} P(A^c) \cdot P(B^c) &= (1 - P(A)) \cdot (1 - P(B)) \\ &= 1 - P(A) - P(B) + P(A) \cdot P(B) \\ &= 1 - (P(A) + P(B) - P(AB)) \\ &= P(A^c B^c) \quad \blacksquare \end{aligned}$$

Exercise 2 (Conditional Probability with Cards)

(Based on **LW** 1.12) Suppose we have three cards:

- The first is green on both sides.
- The second is red on both sides.

- The third is green on one side and red on the other.

We choose a card at random and we see one side. (The side we see is also random.) If the side we see is green, what is the probability that the other side is also green?

Many people intuitively answer 1/2. Calculate the correct answer.

Solution

We simply apply the definition of conditional probability.

$$P(\text{2nd side is green} \mid \text{1st side is green}) = \frac{P(\text{both sides are green})}{P(\text{1st side is green})} = \frac{1/3}{1/3 + 1/3 \cdot 1/2} = \boxed{\frac{2}{3}}$$

Exercise 3 (Discrete Distributions)

Given:

- $X \sim \text{Pois}(\lambda = 3.2)$
- $Y \sim \text{Binom}(n = 20, p = 0.3)$
- $Z \sim \text{Geom}(p = 0.4)$

Here, we assume the probability mass function of Z is given by

$$f(z) = (1-p)^{z-1}p, \quad z = 1, 2, 3, \ldots$$

Note: This is *not* the same parameterization of the geometric distribution that R uses by default.

Calculate:

- $P[1 \leq X \leq 4]$.
- $P[Y > 4]$.
- $P[Z > 3]$.

Solution

$$P[1 \leq X \leq 4] = \boxed{0.74}$$

```
sum(dpois(1:4, lambda = 3.2))
ppois(4, lambda = 3.2) - ppois(0, lambda = 3.2)
```

```
## [1] 0.7398503
## [1] 0.7398503
```

$$P[Y > 4] = \boxed{0.762}$$

```
1 - pbinom(4, size = 20, prob = 0.3)
1 - sum(dbinom(0:4, size = 20, prob = 0.3))
```

```
## [1] 0.7624922
## [1] 0.7624922
```

$$P[Z > 3] = \boxed{0.216}$$

```
1 - pgeom(2, prob = 0.4)
```

```
## [1] 0.216
```

Note that in R, we are calculating $P(W > 2)$ where $W = Z - 1$. Thus,

$$\begin{aligned} P(Z > 3) &= P(Z = 4) + P(Z = 5) + P(Z = 6) + \ldots \\ P(W > 2) &= P(W = 3) + P(W = 4) + P(Z = 5) + \ldots \\ P(Z > 3) &= P(W > 2) \end{aligned}$$

Exercise 4 (Bayes' Theorem)

Given:

- $\text{P}(Y = A) = 0.09$
- $\text{P}(Y = B) = 0.18$
- $\text{P}(Y = C) = 0.73$
- $X \mid Y = A \sim \text{Poisson}(\lambda_A = 4.25)$
- $X \mid Y = B \sim \text{Poisson}(\lambda_B = 6.14)$
- $X \mid Y = C \sim \text{Poisson}(\lambda_C = 8.51)$

Calculate $\text{P}(Y = B \mid X = 3)$.

Solution

For simplicity of notation we define

- $\pi_A = \text{P}(Y = A) = 0.09$
- $\pi_B = \text{P}(Y = B) = 0.18$
- $\pi_C = \text{P}(Y = C) = 0.73$

Also, let $f_k(x)$ be probability mass function of a Poisson distribution with λ_k.

We then apply Bayes' theorem.

$$\begin{aligned} \text{P}(Y = B \mid X = 3) &= \frac{\text{P}(Y = B \cap X = 3)}{\text{P}(X = 3)} \\ &= \frac{\pi_B \cdot f_B(3)}{\pi_A \cdot f_A(3) + \pi_B \cdot f_B(3) + \pi_C \cdot f_C(3)} \end{aligned}$$

Then, we move to R to carry out the calculation.

```
pi = c(0.09, 0.18, 0.73)
pmf = dpois(x = 3, lambda = c(4.25, 6.14, 8.51))
pi * pmf / sum(pi * pmf)
```

```
## [1] 0.3532709 0.3218545 0.3248745
```

So we see that the final answer is

$$\text{P}(Y = B \mid X = 3) = \boxed{0.3218545}.$$

Exercise 5 (Normal Distribution)

(Based on **LW** 2.18) Let $X \sim \text{Normal}(\mu = 2.5, \sigma = 3.2)$.

- Calculate $P(X < 4)$.
- Calculate $P(X > 2)$.
- Find x such that $P(X > x) = 0.05$.
- Calculate $P(X = 4)$.
- Calculate $P(0 < X \leq 3)$.

Solution

We immediately move to R and directly calculate each of the above.

```
pnorm(4, mean = 2.5, sd = 3.2)
```

```
## [1] 0.6803758
```

```
pnorm(2, mean = 2.5, sd = 3.2, lower.tail = FALSE)
```

```
## [1] 0.562082
```

```
qnorm(0.05, mean = 2.5, sd = 3.2, lower.tail = FALSE)
```

```
## [1] 7.763532
```

```
pnorm(4, mean = 2.5, sd = 3.2) - pnorm(4, mean = 2.5, sd = 3.2)
```

```
## [1] 0
```

```
diff(pnorm(c(0, 3), mean = 2.5, sd = 3.2))
```

```
## [1] 0.3447543
```

Exercise 6 (Single Variable Transformation)

(Based on **LW** 2.4) Let X have the probability density function

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

Define $Y = 1/X$. Find the probability density function of Y.

Solution

We first note that $r(x) = 1/x$ is a strictly monotone decreasing function on the support. Then, we can use equation 2.12 from **LW** which gives

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

where $s(y) = 1/y$ is the inverse of r. We then have

$$f_Y(y) = f_X\left(\frac{1}{y}\right) \left| \frac{-1}{y^2} \right| = \begin{cases} \frac{3}{8y^2} & \frac{1}{5} < y < \frac{1}{3} \\ \frac{1}{4y^2} & y > 1 \\ 0 & \text{otherwise} \end{cases}$$

Exercise 7 (Conditional Distribution)

(Based on **LW** 2.17) Given

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

Find $P(X < 0.5 \mid Y = 0.5)$.

Solution

First, we find the marginal distribution of Y.

$$f_Y(y) = \int_0^1 c(x+y)^2 dx = c(y^2 + y + 1/3) \quad 0 < y < 1$$

Then, when $0 < y < 1$ we find the conditional distribution of Y given X.

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{(x+y)^2}{(y^2+y+1/3)} \quad 0 < x < 1$$

Finally, we calculate the particular probability.

$$P\left(X < 0.5 \mid Y = 0.5\right) = \int_0^{0.5} \frac{(x+0.5)^2}{(0.5^2+0.5+1/3)} dx = \boxed{\frac{7}{26}}$$

Exercise 8 (Conditional Poissons)

(Based on **LW** 2.16) Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ and assume that X and Y are independent. Find the distribution of X given that $X + Y = n$.

Solution

First, note that the book provides the hint that $X + Y \sim \text{Poisson}(\lambda + \mu)$.

Also, note that

$$P(X = k, X + Y = n) = P(X = k, Y = n - k) = P(X = k) \cdot P(Y = n - k)$$

Now, we have

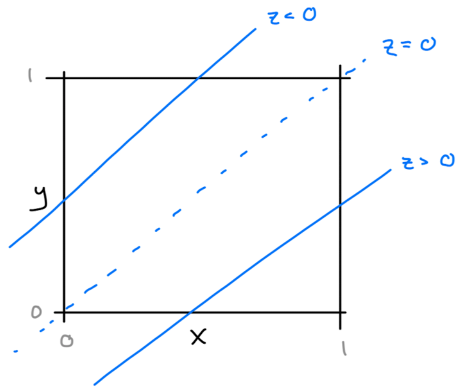
$$\begin{aligned} P(X = k \mid X + Y = n) &= \frac{P(X = k) \cdot P(Y = n - k)}{P(X + Y = n)} \\ &= \frac{\frac{\lambda^k e^{-\lambda}}{k!} \cdot \frac{\mu^{n-k} e^{-\mu}}{(n-k)!}}{\frac{(\lambda+\mu)^n e^{-(\lambda+\mu)}}{n!}} \\ &= \text{... some algebra happens here ...} \\ &= \binom{n}{k} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(1 - \frac{\lambda}{\lambda+\mu}\right)^{n-k} \end{aligned}$$

Thus, X given that $X + Y = n$ has a Binomial $\left(n, \frac{\lambda}{\lambda+\mu}\right)$ distribution.

Exercise 9 (Difference Distribution)

Let $X \sim \text{Uniform}(0,1)$ and $Y \sim \text{Uniform}(0,1)$ be independent. Find the probability density function of $X - Y$.

Solution



Define $Z = X - Y$.

First, we find the CDF.

$$F_Z(z) = P(X - Y < z) = \begin{cases} 0 & z \leq -1 \\ \int_0^{z+1} \int_{x-z}^{1} dy dx = \frac{(z+1)^2}{2} & -1 < z \leq 0 \\ 1 - \int_z^1 \int_0^{x-z} dy dx = 1 - \frac{(1-z)^2}{2} & 0 < z < 1 \\ 1 & z \geq 1 \end{cases}$$

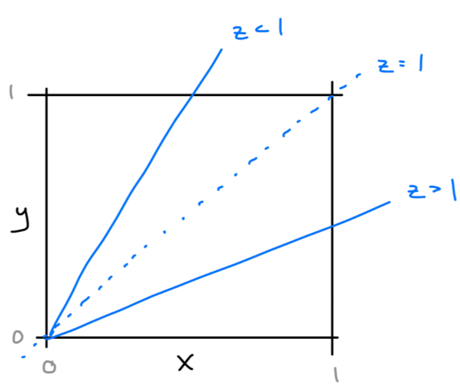
We then obtain the PDF.

$$f_Z(z) = \begin{cases} 1 + z & -1 < z \leq 0 \\ 1 - z & 0 < z < 1 \\ 0 & \text{otherwise} \end{cases}$$

Exercise 10 (Ratio Distribution)

Let $X \sim \text{Uniform}(0, 1)$ and $Y \sim \text{Uniform}(0, 1)$ be independent. Find the probability density function of X/Y.

Solution



Define $Z = X/Y$.

First, we find the CDF.

$$F_Z(z) = P(X/Y < z) = \begin{cases} 0 & z \le 0 \\ \int_0^z \int_{x/z}^1 dy dx = \frac{z}{2} & 0 < z < 1 \\ 1 - \int_0^1 \int_0^{x/z} dy dx = 1 - \frac{1}{2z} & z \ge 1 \end{cases}$$

We then obtain the PDF.

$$f_Z(z) = \begin{cases} \frac{1}{2} & 0 < z < 1 \\ \frac{1}{2z^2} & z \ge 1 \\ 0 & \text{otherwise} \end{cases}$$

Exercise 11 (Simulation Study)

Use a computer experiment to verify your results to Exercise 9 and Exercise 10. For each, do the following:

- Generate a vector $x = (x_1, x_2, \ldots x_{1000})$ from a Uniform$(0, 1)$ distribution.
- Generate a vector $y = (y_1, y_2, \ldots y_{1000})$ from a Uniform$(0, 1)$ distribution.
- Define $z = x - y$ or $z = x/y$.
- Plot a histogram of z. (Be sure to use density histogram, not frequency.)
- Overlay the true density that you calculated.

You may use any computational tool of your choice. Where possible, please supply your code, and of course, the two histograms. If you are using `R`, we recommend using `breaks = 1000`, `xlim = c(0, 25)`, and `ylim = c(0, 0.55)` for the plot of $z = x/y$. (Otherwise, it may be difficult to see that your density matches the histogram.)

Solution

```
# simulate vectors for x and y
set.seed(42)
x = runif(1000)
y = runif(1000)
```

```
# density for z = x - y
f_diff = function (z) {
  ifelse(z < 0, 1 + z, 1 - z)
}
```

```
# density for z = x / y
f_ratio = function(z) {
  ifelse(z < 1, 0.5, 1 / (2 * z ^ 2))
}
```

```
# setup plot structure
par(mfrow = c(1, 2))

# histogram for difference
hist(x - y, probability = TRUE)
box()
grid()
curve(f_diff(x), add = TRUE, col = "red")

# histogram for ratio
hist(x / y, probability = TRUE, breaks = 1000,
     xlim = c(0, 25), ylim = c(0, 0.55))
box()
```

```
grid()
curve(f_ratio(x), add = TRUE, col = "red")
```