51/54 A+

Wu_Youzhi_lab_1

February 19, 2019

1 Lab 1: Probability Theory

1.1 W203: Statistics for Data Science

Student Name: Youzhi Wu
Section Number: 05

1.2 1. Meanwhile, at the Unfair Coin Factory...

You are given a bucket that contains 100 coins. 99 of these are fair coins, but one of them is a trick coin that always comes up heads. You select one coin from this bucket at random. Let T be the event that you select the trick coin. This means that $P(T) = 0.01$.

a. Suppose you flip the coin once and it comes up heads. Call this event H_1. If this event occurs, what is the conditional probability that you have the trick coin? In other words, what is $P(T|H_1)$?

b. Suppose instead that you flip the coin k times. Let H_k be the event that the coin comes up heads all k times. If you see this occur, what is the conditional probability that you have the trick coin? In other words, what is $P(T|H_k)$.

c. How many heads in a row would you need to observe in order for the conditional probability that you have the trick coin to be higher than 99%?

a. what is $P(T|H_1)$

3/3

$$P(T|H_1) = \frac{P(T \cap H_1)}{P(H_1)}$$

$\leftarrow$ $P(T \cap H_1)$ equals to $P(T)$ because the trick coin will always come up heads.
$\leftarrow$ $P(H_1)$ equals to (1) either flipping a fair coin to get heads, (2) or flipping the trick coin to get he

$$= \frac{0.01}{0.99 * 0.5 + 0.01}$$
$$= 0.0198$$

✓

b. what is $P(T|H_k)$

3/3

$$P(T|H_k) = \frac{P(T \cap H_k)}{P(H_k)}$$

$\leftarrow P(T \cap H_k)$ equals to the possibility of flipping trick coin which would be $P(T)$.
$\leftarrow P(H_k)$ equals to the possibility of getting event H_1 k times, which would be $P(H_1)^k$.

$$= \frac{0.01}{0.99 \cdot 0.5^k + 0.01}$$

3/3

c. what is the k in order for $P(T|H_k)$ to be higher than 99%

Because we have $P(T|H_k) = \frac{0.01}{0.99 \cdot 0.5^k + 0.01}$ which needs to be higher than 99%, therefore

$$\frac{0.01}{0.99 \cdot 0.5^k + 0.01} > 0.99$$
$$k > 2\log_2 99$$
$$k > 13.259$$

Therefore, you need to observe at least 14 times heads in a row in order for the conditional probability that you have the trick coin to be higher than 99%.

1.3 2. Wise Investments

You invest in two startup companies focused on data science. Thanks to your growing expertise in this area, each company will reach unicorn status (valued at \$1 billion) with probability 3/4, independent of the other company. Let random variable X be the total number of companies that reach unicorn status. X can take on the values 0, 1, and 2. Note: X is what we call a binomial random variable with parameters $n = 2$ and $p = 3/4$.

a. Give a complete expression for the probability mass function of X.

b. Give a complete expression for the cumulative probability function of X.

c. Compute $E(X)$.

d. Compute $var(X)$.

3/3

a. Give a complete expression for the probability mass function of X.

Because the distribution follows binomial distribution with parameters $n = 2$ and $p = 3/4$, the pmf of X, $P(X)$, is:

$$P(X) = \begin{cases} \frac{1}{16} & \text{if } X = 0 \\ \frac{3}{8} & \text{if } X = 1 \\ \frac{9}{16} & \text{if } X = 2 \\ 0 & \text{if otherwise} \end{cases}$$

3/3

b. Give a complete expression for the cumulative probability function of X.

The cmf of X, $B(X)$, is:

$$B(X) = \begin{cases} \frac{1}{16} & \text{if } X \leq 0 \\ \frac{7}{16} & \text{if } X \leq 1 \\ 1 & \text{if } X \leq 2 \\ 1 & \text{if } x > 2 \end{cases}$$

3/3

c. Compute $E(X)$.

$$\begin{aligned} E(X) &= \sum_{x=0}^{x=2} x \cdot P(X = x) \\ &= 0 \cdot \frac{1}{16} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{9}{16} \\ &= 0 + \frac{3}{8} + \frac{9}{8} \\ &= \frac{3}{2} \end{aligned}$$

3/3

d. Compute $var(X)$.

$$\begin{aligned} var(X) &= E(X^2) - [E(X)]^2 \\ &= \sum_{x=0}^{x=2} x^2 \cdot P(X = x) - [E(X)]^2 \\ &= 0^2 \cdot \frac{1}{16} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{9}{16} - \left[\frac{3}{2}\right]^2 \\ &= 0 + \frac{3}{8} + \frac{9}{4} - \frac{9}{4} \\ &= \frac{3}{8} \end{aligned}$$

1.4 3. A Really Bad Darts Player

Let X and Y be independent uniform random variables on the interval $[-1, 1]$. Let D be a random variable that indicates if (X, Y) falls within the unit circle centered at the origin. We can define D as follows:

$$D = \begin{cases} 1, & X^2 + Y^2 < 1 \\ 0, & otherwise \end{cases}$$

Note that D is a Bernoulli variable.

a. Compute the expectation $E(D)$. Hint: it might help to remember why we use area diagrams to represent probabilites.

b. Compute the standard deviation of D.

c. Write an R function to compute the value of D, given a value for X and a value for Y. Use R to simulate a draw for X and a draw for Y, then compute the value of D.

d. Use R to simulate the previous experiment 1000 times, resulting in 1000 samples for D. Compute the sample mean and sample standard deviation of your result, and compare them to the true values in parts a. and b.

3/3

a. Compute the expectation $E(D)$.

The total area of the sample space is $(1+1)^2 = 4$ since X and Y follows uniform distribution on the interval of $[-1,1]$. Then the pmf of D, $P(D)$, is:

$$P(D) = \begin{cases} \frac{\pi \cdot 1^2}{4} & \text{if } D = 1 \\ 1 - \frac{\pi \cdot 1^2}{4} & \text{if } D = 0 \\ 0 & \text{if otherwise} \end{cases}$$

Therefore, $P(D)$ could be written as follows:

$$P(D) = \begin{cases} \frac{\pi}{4} & \text{if } D = 1 \\ 1 - \frac{\pi}{4} & \text{if } D = 0 \\ 0 & \text{if otherwise} \end{cases}$$

Therefore, $E(D)$ can be computed as follows:

$$\begin{aligned} E(D) &= \sum_{d=0}^{d=1} d \cdot P(D=d) \\ &= 0 \cdot (1 - \frac{\pi}{4}) + 1 \cdot \frac{\pi}{4} \\ &= \frac{\pi}{4} \\ &\approx 0.785 \end{aligned}$$

b. Compute the standard deviation of D.

3/3

The standard deviation of D, $S(D)$ equals to the positive square root of the $var(D)$. The $var(D)$ equals to $E(D^2) - [E(D)]^2$. Therefore, the $S(D)$ can be computed as follows:

$$\begin{aligned} S(D) &= \sqrt{var(D)} \\ &= \sqrt{E(D^2) - [E(D)]^2} \\ &= \sqrt{\sum_{d=0}^{d=1} d^2 \cdot P(D=d) - [E(D)]^2} \\ &= \sqrt{0^2 \cdot (1 - \frac{\pi}{4}) + 1^2 \cdot \frac{\pi}{4} - \left[\frac{\pi}{4}\right]^2} \\ &= \frac{\sqrt{4\pi - \pi^2}}{4} \\ &\approx 0.411 \end{aligned}$$

2/3

c. Write an R function to compute the value of D, given a value for X and a value for Y. Use R to simulate a draw for X and a draw for Y, then compute the value of D.

not a function

```
In [2]: # set seed
        set.seed(89)

        # Take Draws
        n = 1
        x <- runif(n, min = -1, max = 1)
        y <- runif(n, min = -1, max = 1)

        # compute the value of D
        if((x^2 + y^2) > 1) {
            d <- 1
        } else {
            d <- 0
        }
        print(d)

[1] 1
```

2/3

d. Use R to simulate the previous experiment 1000 times, resulting in 1000 samples for D. Compute the sample mean and sample standard deviation of your result, and compare them to the true values in parts a. and b.

```
In [5]: # set seed
        set.seed(89)

        reps = 0
        d_vector <- c()

        repeat{
            # Take Draws
            n = 1
            x <- runif(n, min = -1, max = 1)
            y <- runif(n, min = -1, max = 1)

            # compute the value of D
            if((x^2 + y^2) > 1) {
                d <- 1
            } else {
                d <- 0
            }

            # store the value of D in a vector
            d_vector <- c(d_vector, d)

            # count the # of experiments, if over 1000 times, break from repeat
```

```
        reps = reps + 1
        if(reps > 1000) {
            break
        }
    }

    mean_obs <- mean(d_vector)
    sd_obs <- sd(d_vector)
    print(mean_obs)
    print(sd_obs)
```

```
[1] 0.2227772
[1] 0.416318
```

Correct.

2 4. Relating Min and Max

Continuous random variables X and Y have a joint distribution with probability density function,

$$f(x,y) = \begin{cases} 2, & 0 < y < x < 1 \\ 0, & otherwise. \end{cases}$$

You may wonder where you would find such a distribution. In fact, if A_1 and A_2 are independent random variables uniformly distributed on $[0,1]$, and you define $X = max(A_1, A_2)$, $Y = min(A_1, A_2)$, then X and Y will have exactly the joint distribution defined above.

a. Draw a graph of the region for which X and Y have positive probability density.

b. Derive the marginal probability density function of X, $f_X(x)$. Make sure you write down a complete expression.

c. Derive the unconditional expectation of X.

d. Derive the conditional probability density function of Y, conditional on X, $f_{Y|X}(y|x)$

e. Derive the conditional expectation of Y, conditional on X, $E(Y|X)$.

f. Derive $E(XY)$. Hint 1: Use the law of iterated expectations. Hint 2: If you take an expectation conditional on X, X is just a constant inside the expectation. This means that $E(XY|X) = XE(Y|X)$.

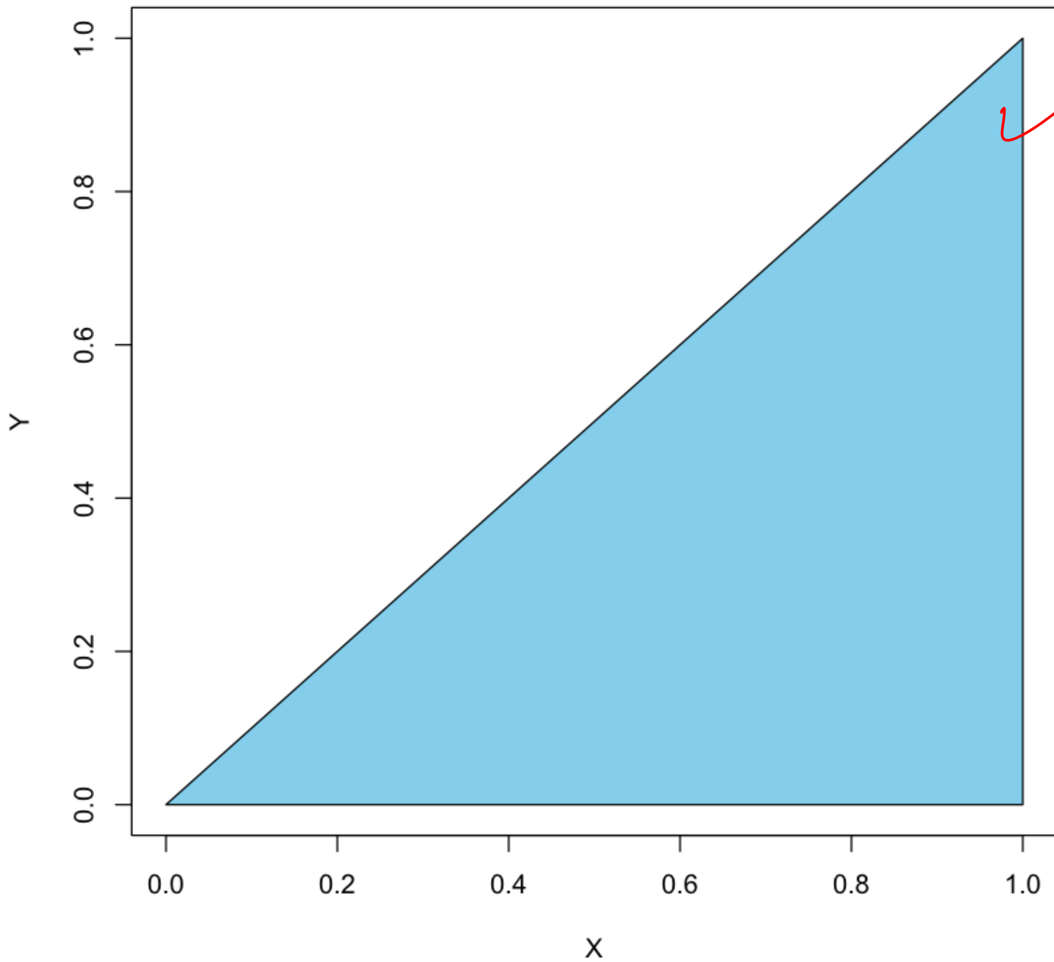
g. Using the previous parts, derive $cov(X,Y)$

a. Draw a graph of the region for which X and Y have positive probability density.

3/3

The graph is shown as below. The skyblue region is where X and Y have postive prbability density.

```
In [9]: x <- c(0, 1, 1) # The x-coordinate of the vertices
        y <- c(0, 1, 0) # The y-coordinate of the vertices
        plot(c(0,1), c(0,1), type = "n", xlab = "X", ylab = "Y")
        polygon(x, y, col = 'skyblue')
```



3/3

b. Derive the marginal probability density function of X, $f_X(x)$.

The marginal probability density function of X can be written as below:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\ \forall x \in (0,1) \qquad &= \int_0^x 2dy \\ &= [2y]_0^x \\ &= 2x \end{aligned}$$

> Therefore, we have

$$f_X(x) = \begin{cases} 2x & \text{if } x \in (0,1) \\ 0 & \text{if otherwise} \end{cases}$$

3/3

c. Derive the unconditional expectation of X.

The the unconditional expectation of X can be computed as below:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x \cdot f_X(x)dx \\
&= \int_0^1 x \cdot 2xdx \\
&= \left[\frac{2x^3}{3}\right]_0^1 \\
&= \frac{2}{3}
\end{aligned}$$

2/3

d. Derive the conditional probability density function of Y, conditional on X, $f_{Y|X}(y|x)$.

The the conditional probability density function can be computed as follows:

$$\begin{aligned}
f_{Y|X}(y|x) &= \frac{f(x,y)}{f_X(x)} \\
&= \frac{2}{2x} \\
&= \frac{1}{x} \quad \text{for } y \in (0, x)
\end{aligned}$$

3/3

e. Derive the conditional expectation of Y, conditional on X, $E(Y|X)$.

The conditional expectation can be computed as follows:

$$\begin{aligned}
E_Y(Y|X) &= \int_Y y \cdot f_{Y|X}(y|x)dy \\
&= \int_0^x y \cdot \frac{1}{x}dy \\
&= \frac{1}{x} \cdot \int_0^x ydy \\
&= \frac{1}{x} \cdot \left[\frac{y^2}{2}\right]_0^x \\
&= \frac{1}{x} \cdot \frac{x^2}{2} \\
&= \frac{x}{2}
\end{aligned}$$

f. Derive $E(XY)$.

The expectation can be computed as below:

$$
\begin{aligned}
E(XY) &= E_X[E_Y(XY|X)] \\
&= E_X[X \cdot E_Y(Y|X)] \\
&= E_X[X \cdot \frac{X}{2}] \\
&= E_X[\frac{X^2}{2}] \\
&= \int_0^1 \frac{x^2}{2} \cdot 2x dx \\
&= \left[\frac{x^4}{4}\right]_0^1 \\
&= \frac{1}{4}
\end{aligned}
$$

3/3

g. Using the previous parts, derive $cov(X, Y)$

$$
\begin{aligned}
cov(X,Y) &= E(XY) - E(X)E(Y) \\
&= E(XY) - E(X)E_X[E_Y(Y|X)] \\
&= E(XY) - E(X) \cdot E_X[\frac{X}{2}] \\
&= E(XY) - E(X) \cdot \int_0^1 \frac{x}{2} \cdot 2x dx \\
&= E(XY) - E(X) \cdot \left[\frac{x^3}{3}\right]_0^1 \\
&= \frac{1}{4} - \frac{2}{3} \cdot \frac{1}{3} \\
&= \frac{1}{36}
\end{aligned}
$$

```
In [ ]:
```