

Lab 1: Probability Theory

49/54

A

W203: Statistics for Data Science

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 $P(T) = 0.01$.

- 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 $P(T|H_1)$?
- 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 $P(T|H_k)$.
- 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%?

- Known probabilities:**

$$P(T) = 0.01$$

$$P(!T) = 0.99$$

$$P(H|T) = 1$$

$$P(H|!T) = 0.5$$

2/3

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 $P(T|H_1)$?

- The probability of flipping the trick coin, given you flipped a heads $P(T|H_1)$:

$$\begin{aligned}
 P(T|H_1) &= \frac{P(T \cap H_1)}{P(H_1)} \\
 &= \frac{P(H_1 \cap T)}{P(H_1)} \\
 &= \frac{P(H_1|T)P(T)}{P(H_1)} \\
 &= \frac{(1)(0.01)}{\cancel{0.05} \text{ :505}} \quad \text{split into } P(H_1|T) + P(H_1|\neg T) \text{ see solutions} \\
 &= \frac{0.01}{0.50} \\
 &= 0.02 \text{ (y.)}
 \end{aligned}$$

★ The conditional probability of flipping the trick coin, given you flipped a head, $P(T|H_1) = 0.02$.

3/3 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 $P(T|H_k)$.

- The probability of observing k heads, given that the coin is unfair, T , is $P(T|H_k)$:

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

- To calculate this probability, we first need $P(H_k)$

- The probability of getting k heads is $P(H_k)$:

$$\begin{aligned} P(H_k) &= (P(!T) \cdot P(H_k|!T)) + (P(T) \cdot P(H_k|T)) \\ &= (P(!T) \cdot P(H_k|!T)^k) + (P(T) \cdot P(H|T)^k) \\ &= (0.99 \cdot 0.5^k) + (0.01 \cdot 1^k) \\ &= (0.99 \cdot 0.5^k) + 0.01 \end{aligned}$$

- Since we now know $P(H_K)$, we need $P(H_k|T)$:

$$P(H_k|T) = P(H_k|T) = P(H|T)^k = 1^k = 1$$

- Now, to solve $P(T|H_k)$:

$$\begin{aligned} P(T|H_k) &= \frac{P(H_k|T) \cdot P(T)}{P(H_k)} \\ &= \frac{1 \cdot 0.01}{(0.99 \cdot 0.5^k) + 0.01} \\ &= \frac{0.01}{(0.99 \cdot 0.5^k) + 0.01} \end{aligned}$$

★ Thus, $P(T|H_k) = \frac{0.01}{(0.99 \cdot 0.5^k) + 0.01}$ ✓

c.

2/3

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%?

- The number of k heads in a row are needed for $P(T|H_k) > 0.99$:

$$\frac{0.01}{(0.99 \cdot 0.5^k) + 0.01} > 0.99$$

$$(0.99 \cdot 0.5^k) + 0.01 \cdot 0.01 > 0.99 \cdot (0.99 \cdot 0.5^k) + 0.01$$

$$\frac{0.01}{0.99} > \frac{0.99 \cdot (0.99 \cdot 0.5^k) + 0.01}{0.99}$$

$$\frac{0.01}{0.99} > (0.99 \cdot 0.5^k) + 0.01$$

$$(0.99 \cdot 0.5^k) < \frac{1}{99} - 0.01$$

$$(0.99 \cdot 0.5^k) < \frac{1}{99} - \frac{1}{100}$$

$$(0.99 \cdot 0.5^k) < \frac{100}{9900} - \frac{99}{9900}$$

$$(0.99 \cdot 0.5^k) < \frac{1}{9900}$$

$$0.5^k < \frac{1}{(9900 \cdot 0.99)}$$

$$0.5^k < \frac{1}{9801}$$

★ The number of k heads in a row needed to inflate

$$P(T|H_k) > 0.99 = 0.5^k < \frac{1}{9801} \quad \text{— solve for } k$$

(✓)

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 $p = 3/4$.

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

c. Compute $E(X)$.

d. Compute $\text{var}(X)$.

a. 3/3

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

- The $p(x)$ for each $x = 0, 1, 2$ can be calculated from the binomial distribution pmf:

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

- where

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

- For $x = 0, 1, 2$:

$$p(X = 0):$$

$$\begin{aligned} p(0) &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= \frac{2!}{0!(2-0)!} \left(\frac{3}{4}\right)^0 \left(\frac{1}{4}\right)^{2-0} \\ &= \frac{2}{2} \cdot 1 \cdot 0.0625 \\ &= 1 \cdot 1 \cdot 0.0625 \\ &= 0.0625 \\ &= 0.0625 \end{aligned}$$

$$p(X = 1):$$

$$\begin{aligned} p(1) &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\ &= \frac{2!}{1!(2-1)!} \left(\frac{3}{4}\right)^1 \left(\frac{1}{4}\right)^{2-1} \\ &= \frac{2}{1} \cdot 0.75 \cdot 0.25 \\ &= 2 \cdot 0.75 \cdot 0.25 \\ &= 0.0625 \\ &= 0.375 \end{aligned}$$

$$p(X = 2):$$

$$\begin{aligned}
 p(2) &= \binom{n}{x} p^x (1-p)^{n-x} \\
 &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
 &= \frac{2!}{2!(2-2)!} \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^{2-2} \\
 &= \frac{2}{2} \cdot 0.5625 \cdot 1 \\
 &= 1 \cdot 0.5625 \cdot 1 \\
 &= 0.5625
 \end{aligned}$$

★ In summary, the probability mass function is:

x	0	1	2
P(X=x)	0.0625	0.375	0.5625

$$p(X = x) = \begin{cases} 0.0625 & x = 0 \\ 0.375 & x = 1 \\ 0.5625 & x = 2 \end{cases}$$

3/3

b.

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

★ Case by case:

$$F(0) = P(x \leq 0) = P(x = 0) = p(0) = 0.0625$$

$$F(1) = P(0 < x \leq 1) = P(x = 0 \text{ or } 1) = p(0) + p(1) = 0.0625 + 0.375 = 0.4375$$

$$F(2) = P(1 < x \leq 2) = P(x = 0 \text{ or } 1 \text{ or } 2) = p(0) + p(1) + p(2) = 0.0625 + 0.375 + 0.5625 = 1$$

★ For each case:

$$F(x) = \begin{cases} 0.0625 & x \leq 0 \\ 0.4375 & 0 < x \leq 1 \\ 1 & 1 < x \leq 2 \end{cases} \quad \checkmark$$

c.

3/3

Compute $E(X)$

- For discrete random variables:

$$\begin{aligned} E(X) &= E(f(x)) = \sum f(x)p(x) \\ &= [(0 \cdot 0.625) + (1 \cdot 0.375) + (2 \cdot 0.5625)] \\ &= 1.5 \quad \checkmark \end{aligned}$$

★ In the long run, an expected 1.5 companies will reach unicorn status.

3/3 d.

Compute $\text{var}(X)$.

- Since:

$$\begin{aligned} \text{Var}(X) &= E(x^2) - [E(x)]^2 \\ &= (x^2) - [1.5]^2 \end{aligned}$$

$$\text{where } E(x^2) = [(0^2 \cdot 0.625) + (1^2 \cdot 0.375) + (2^2 \cdot 0.5625)]$$

$$= [(0) + (0.375) + (2.25)]$$

$$= [(0) + (0.375) + (2.25)]$$

$$= 2.625$$

$$\text{So... } Var(X) = E(x^2) - [E(x)]^2$$

$$= (2.625) - [1.5]^2$$

$$= 2.625 - 2.25$$

$$= 0.375$$

★ The $var(x) = 0.375$. ✓

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:

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

Note that D is a Bernoulli variable.

- Compute the expectation $E(D)$. Hint: it might help to remember why we use area diagrams to represent probabilities.
- Compute the standard deviation of D .
- 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 .

3/3

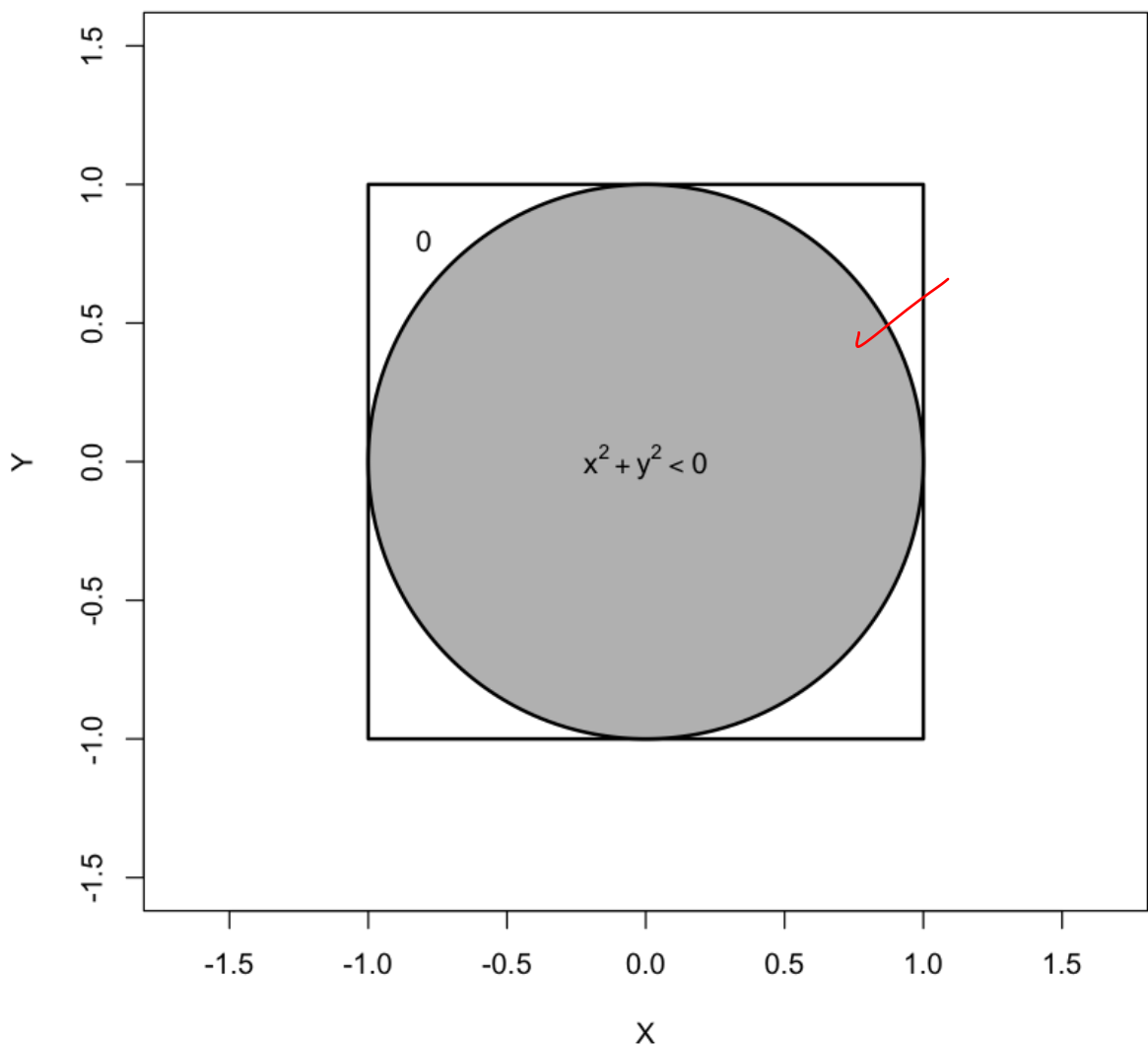
a.

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

- The $E(D)$ can be derived from an area diagram.

```
In [3]: # Install package
install.packages("plotrix")
# Invoke package
library("plotrix")
# Create empty plot
plot(NULL,xlim=c(-1.5,1.5), ylim=c(-1.5, 1.5), col = "white",
      xlab = "X", ylab = "Y", asp=1)
# Plot square
polygon(x = c(-1,1,1,-1),
        y = c(-1,-1,1,1),
        lwd = 2)
# Plot function
draw.circle(0,0,1,nv=1000,border=NULL,col="grey",lty=1,lwd=2)
# Name the function inside the circle
text(0, 0, expression(x^{2} + y^{2} < 0))
# Name the function inside the circle
text(-0.8, 0.8, expression(0))
```

Updating HTML index of packages in '.Library'
Making 'packages.html' ... done



The expected value of D , $E[D]$:

$$E[D] = E(f(d)) = \sum f(d)p(d)$$

$$= \left[\left(0 \cdot \frac{1-\pi}{4} \right) + \left(1 \cdot \frac{\pi}{4} \right) \right]$$

$$= \frac{\pi}{4} \checkmark$$

★ Thus, the expected value of D , $E[D] = \frac{\pi}{4}$

b. ~~1/3~~

Compute the standard deviation of D .

• From: *o/s not sample standard dev. compute the theoretical (true) standard deviation*

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

$$= \sqrt{\left[\frac{\left(0 - \frac{\pi}{4}\right)^2 + \left(1 - \frac{\pi}{4}\right)^2}{2-1} \right]}$$

$$= 0.814189304364331$$

★ Thus, the standard deviation of $D = 0.814189304364331$ *(o/c)*

```
In [31]: sqrt(((0-(pi/4))^2 + (1-(pi/4))^2)/(2-1))
```

0.814189304364331

c. *3/5*

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 .

```
In [4]: # Function to compute value of D, given x and y
D_value = function (x,y){
  if (x**2 + y**2 < 1.0){
    1
  }
  else{
    0
  }
}
```

```
In [5]: # Draw values of x and y
set.seed(123)
x = runif(1, min = -1, max = 1)
y = runif(1, min = -1, max = 1)
```

```
In [6]: # Function output
D_value(x,y)
```

1

3/3

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 [32]: # Function to calculate E(D): ✓
EX_D = function (n){
  # Initialize mean
  Mean = 0
  # Initialize x
  x = runif(n, min = -1, max = 1)
  # Initialize y
  y = runif(n, min = -1, max = 1)
  # For every value in n
  for (i in seq(n)){
    # Compute the mean as the previous mean plus each x[i], y[i]
    Mean = Mean + D_value(x[i], y[i])
  }
  # Divided by n
  Mean = Mean/n
  # Return the mean
  Mean
}
```

could also just invoke your function from (c)

```
In [33]: EX_D(1000)
```

0.788

★ The expected value of D , $E[D]$ computed without simulation, $E[D] = \frac{\pi}{4}$ is very close to the value computed with simulation where $E[D] = 0.788$. This is to be expected for the expected value, since it represents the expected value in the long run!

```
In [28]: # Function to caculate sd(D):
SD_D = function (n){
  # Initialize mean
  Mean = 0
  # Intialize standard deviation
  Std = 0
  # Initialize x
  x = runif(n, min = -1, max = 1)
  # Initialize y
  y = runif(n, min = -1, max = 1)
  # For every ith value in n
  for (i in seq(n)){
    # Compute the mean as the previous mean plus each x[i], y[i]
    Mean = Mean + D_value(x[i], y[i])
  }
  # Divided by n
  Mean = Mean/n
  # For every jth value in n
  for (j in seq(n)){
    # Calculate the standard deviation
    Std = Std + (D_value(x[j], y[j]) - Mean)^2
  }
  # Square root of standard deviation
  Std = sqrt(Std/(n-1))
  # Return the standard deviation
  Std
}
```

```
In [29]: SD_D(1000)
```

0.409632516774238

★ The standard deviation of D , σ computed without simulation, $\sigma = 0.814$. is not very close to the value computed with simulation where $\sigma = 0.409$. The latter is \sim half the size of the former. This is also to be expected for the standard deviation, since it is more sensitive to n than the expected value.

4. Relating Min and Max

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

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

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.

- Draw a graph of the region for which X and Y have positive probability density.
- 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. 3/3

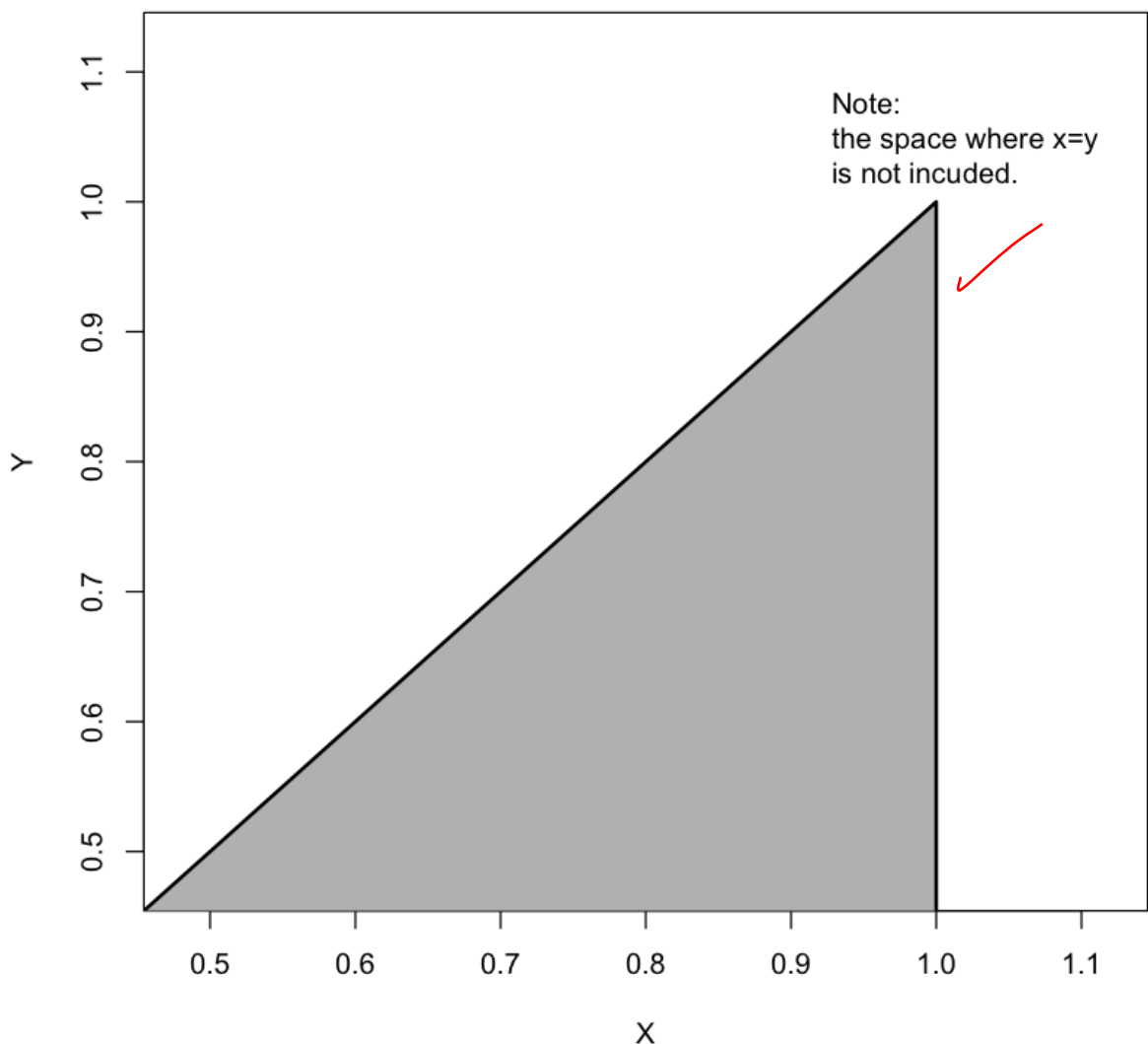
Draw a graph of the region for which X and Y have positive probability density.

- Together, the x and y jointly represent the area that satisfies the constraint where $y < x$, or a right triangular. ✓

```
In [4]: # Create empty plot
plot(0.8, 0.8, col = "white",
      xlab = "X", ylab = "Y")
# Plot function
polygon(x = c(0,1,1),
        y = c(0,1,0),
        col = "grey",
        lwd = 2)
# Name the function
text(1.02, 1.02, expression('Note:\nthe space where x=y\nis not included.'))
```

Warning message in text.default(1.02, 1.02, expression("Note:\nthe space where x=y\nis not included.")):

"font metrics unknown for character 0xa"Warning message in text.default(1.02, 1.02, expression("Note:\nthe space where x=y\nis not included.")):
 "font metrics unknown for character 0xa"Warning message in text.default(1.02, 1.02, expression("Note:\nthe space where x=y\nis not included.")):
 "font metrics unknown for character 0xa"Warning message in text.default(1.02, 1.02, expression("Note:\nthe space where x=y\nis not included.")):
 "font metrics unknown for character 0xa"



b. 2/3

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

- The marginal pdf of X , $f_X(x)$:

- Given that:

$$f_X(x) = \int_{y=0}^x f(x, y) dy$$

$$= 2[y]_0^x$$

$$= 2[x - 0]$$

$$= 2x \quad \text{for } x \in (0, 1)$$

- ★ The marginal pdf of X , $f_X(x)$: $2x$

c. 3/3

Derive the unconditional expectation of X .

- The unconditional expectation of X , $E[X]$:

- Given that:

$$E[X] = E_Y [E_X[X|Y]]$$

$$= E_Y \left[\int_X x \cdot f_{X|Y}(x, y) dx \right]$$

$$= \int_Y \int_X x \cdot f_{X|Y}(x, y) dx f_Y(y) dy$$

$$= \int_Y \int_X x \cdot f_{X|Y}(x, y) f_Y(y) dx dy$$

$$= \int_{x=0}^1 \int_{y=0}^x x f(x, y) dx dy$$

$$= \int_{x=0}^1 \int_{y=0}^x x \cdot 2 dx dy$$

$$= 2 \int_0^1 [y]_0^x dx$$

$$= 2 \left[\frac{x^3}{3} \right]_0^1$$

$$= 2 \left[\frac{1^3}{3} - \frac{0^3}{3} \right]$$

$$= 2 \left[\frac{1}{3} - \frac{0}{3} \right]$$

$$= 2 \left[\frac{1}{3} \right]$$

$$= \frac{2}{3}$$


★ The unconditional expectation of X , $E[X] = \frac{2}{3}$

d. $\frac{3}{3}$

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

- Given that:

$$f_{Y|X}(y|x) = \frac{f_{Y,X}(y|x)}{f_X(x)}, \text{ when } 0 < y < x \quad \checkmark$$

$$= \frac{2}{2x}$$

$$= \frac{1}{x} \quad \checkmark$$

★ The conditional probability density function of Y , conditioned on X $f_{Y|X}(y|x) = \frac{1}{2x}$

e. $\frac{3}{3}$

Derive the conditional expectation of Y , conditional on X , $E[Y|X]$.

- Given that:

$$E[Y|X] = \int_{y=0}^x y f(y|x) dy$$

$$= \int_{y=0}^x y \frac{1}{x} dy$$

$$= \frac{1}{x} \cdot \left[\frac{y^2}{2} \right]_0^x$$

$$= \frac{1}{x} \cdot \left[\frac{x^2}{2} - \frac{0^2}{2} \right]$$

$$= \frac{1}{x} \cdot \left[\frac{x^2}{2} \right]$$

$$= \frac{x}{2} \quad \checkmark$$

★ The conditional expectation of Y , conditioned on X , $E[Y|X] = \frac{x}{2}$

f. ~~3/3~~

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 $E(XY|X) = XE(Y|X)$.

- Given that:

$$\begin{aligned}
 E[XY] &= \int_{x=0}^1 E[XY|X] f_X(x) dx \\
 &= \int_{x=0}^1 x E[Y|X] f_X(x) dx \\
 &= \int_{x=0}^1 x \frac{x}{2} 2x dx \\
 &= \left[\frac{x^4}{4} \right]_0^1 \\
 &= \left[\frac{1^4}{4} - \frac{0^4}{4} \right]_0^1 \\
 &= \frac{1}{4}
 \end{aligned}$$

★ The expectation of YX , $E[YX] = \frac{1}{4}$ ✓

g.

3/3

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

- Given that:

$$\text{cov}[X, Y] \equiv E[XY] - E[X]E[Y]$$

$$= \frac{1}{4} - \left(\frac{2}{3} \cdot E[Y] \right)$$

$$= \frac{1}{4} - \left(\frac{2}{3} \cdot \int_{x=0}^1 E[Y|X] f_X(x) dx \right)$$

$$= \frac{1}{4} - \left(\frac{2}{3} \cdot \int_{x=0}^1 \frac{x}{2} 2x dx \right)$$

$$= \frac{1}{4} - \left(\frac{2}{3} \cdot \left[\frac{x^3}{3} \right]_0^1 \right)$$

$$= \frac{1}{4} - \left(\frac{2}{3} \cdot \left[\frac{1^3}{3} - \frac{0^3}{3} \right] \right)$$

$$= \frac{1}{4} - \left(\frac{2}{3} \cdot \frac{1}{3} \right)$$

$$= \frac{1}{4} - \frac{2}{9}$$

$$= \frac{9}{36} - \frac{8}{36}$$

$$= \frac{1}{36}$$


★ The covariance of X, Y , $\text{cov}[X, Y] = \frac{1}{36}$