MATH 7241 F20

Solutions for Problem Set # 1

MATH 7241: Problems for Notes #1

Due date: Tuesday September 22

Exercise 1 You have two coins, one is unbiased, the other is biased with probability of Heads equal to $2/3$. You toss both coins twice, X is the number of Heads for the fair coin, Y is the number of Heads for the biased coin. Find $P(X > Y)$.

$$\mathbb{P}(X>Y) = \mathbb{P}(X=2, Y=0) + \mathbb{P}(X=2, Y=1) + \mathbb{P}(X=1, Y=0)$$

$$= \mathbb{P}(X=2)\,\mathbb{P}(Y=0) + \mathbb{P}(X=2)\,\mathbb{P}(Y=1) + \mathbb{P}(X=1)\,\mathbb{P}(Y=0)$$

(because X, Y independent)

X	0	1	2
prob	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Y	0	1	2
prob	$(1-p)^2$	$2p(1-p)$	p^2

$p = 2/3$

$$\Rightarrow P(X>Y) = \left(\frac{1}{4}\right)(1-p)^2 + \frac{1}{4}\,2p(1-p)$$

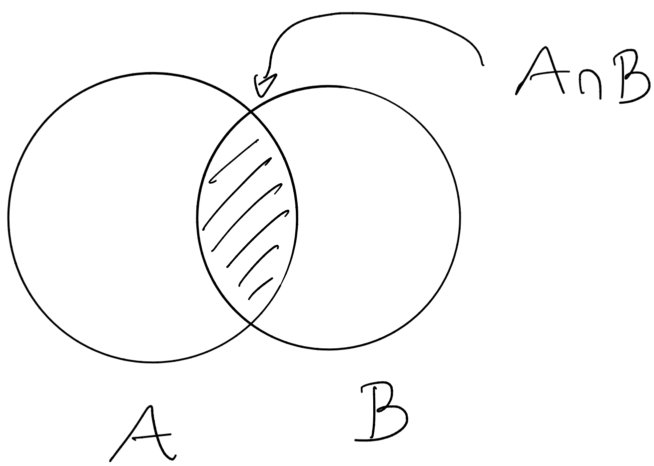
$$+ \frac{1}{2}(1-p)^2$$

$$= \frac{3}{4}\left(\frac{1}{3}\right)^2 + \frac{1}{2}\,\frac{2}{3}\,\frac{1}{3}$$

$$= \frac{7}{36}$$

$\boxed{\textbf{Exercise 2}}$ Let A and B be events such that $P(A) = 0.7$ and $P(B) = 0.9$. Find the largest and smallest possible values of $P(A \cup B) - P(A \cap B)$ (note: the event $A \cup B$ means either A or B or both are true, the event $A \cap B$ means both A and B are true).

$$\mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \cup B) + \mathbb{P}(A \cap B)$$



$$\Rightarrow \mathbb{P}(A \cup B) - \mathbb{P}(A \cap B)$$

$$= 2\,\mathbb{P}(A \cup B) - \mathbb{P}(A) - \mathbb{P}(B)$$

$$= 2\mathbb{P}(A \cup B) - 1.6$$

$\mathbb{P}(A \cup B)$: max. achieved with

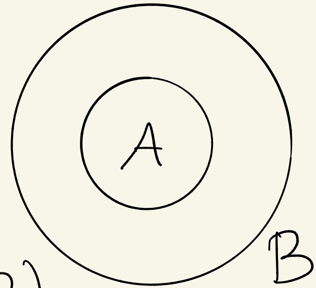
$A \cup B = S$ = whole sample space

$$\Rightarrow \max\left(\mathbb{P}(A \cup B)\right) = 1$$

: min achieved with

$$A \subset B$$

$$\Leftrightarrow \quad A \cup B = B$$



$$\Rightarrow \min\left(\mathbb{P}(A \cup B)\right) = \mathbb{P}(B)$$

$$= 0.9$$

$$\Rightarrow \max\left\{\mathbb{P}(A \cup B) - \mathbb{P}(A \cap B)\right\} = 2 - 1.6 = 0.4$$

$$\min\left\{\mathbb{P}(A \cup B) - \mathbb{P}(A \cap B)\right\} = 1.8 - 1.6 = 0.2$$

Exercise 3 A town has five hotels; three people arrive and each randomly and independently selects a hotel. Find the probability that exactly two of them stay in the same hotel.

All choices are equally likely, so we can compute the probability by counting:

$$\mathbb{P}(\text{two in same hotel}) = \frac{\#\{\text{choices for two in same}\}}{\#\{\text{choices}\}}$$

It is easiest ~~to~~ count by assuming ordered events; so let's assume that the people arrive in order, one after another.

$$\Rightarrow \#\{\text{choices}\} = (5)(5)(5)$$

$\#\{$choices for two in same hotel$\}$

$$= (5)(1)(4)$$ ← first two are same, third is different

$$+ (5)(4)(1)$$ ← first & third same, second different

$$+ (5)(4)(1)$$ ← second and third same, first different

$$= (5)(4)(3)$$

$\Rightarrow \mathbb{P}($two in same hotel$)$

$$= \frac{(5)(4)(3)}{(5)(5)(5)} = \frac{12}{25}$$

$\boxed{\textbf{Exercise 4}}$ Find the mean and variance of the geometric distribution:

$$P(X = n) = (1-p)^{n-1}p, \quad n = 1, 2, \ldots$$

$$\underline{\text{Mean}}: \mathbb{E}[X] = \sum_{n=1}^{\infty} n \cdot \mathbb{P}(X = n)$$

(sum over all possible values)

(same function)

(pdf)

$$\Rightarrow \mathbb{E}[X] = \sum_{n=1}^{\infty} n \cdot (1-p)^{n-1} \cdot p.$$

Here is a trick for doing this kind of sum: recall the formula for a geometric series

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{if } |x| < 1.$$

Take $\frac{d}{dx}$ of both sides:

$$\Rightarrow \sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}$$

Apply to our case: $x = 1-p$

$$\mathbb{E}[X] = \frac{p}{(1-(1-p))^2} = \frac{p}{p^2} = \frac{1}{p}.$$

<u>Varance</u>:

$$VAR[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$\mathbb{E}[X^2] = \sum_{n=1}^{\infty} n^2\, \mathbb{P}(X=n)$$

$$= \sum_{n=1}^{\infty} n^2 (1-p)^{n-1} \cdot p$$

Hee is a variation on our previous trick:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

$$\frac{d}{dx}: \quad \sum_{n=1}^{\infty} n x^{n-1} = \frac{1}{(1-x)^2}$$

$$\frac{d^2}{dx^2}: \quad \sum_{n=2}^{\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}$$

Apply to our problem with

$x = 1-p$:

$$\mathbb{E}[X(X-1)] = \sum_{n=2}^{\infty} n(n-1)(1-p)^{n-1} p$$

$$= p(1-p) \sum_{n=2}^{\infty} n(n-1)(1-p)^{n-2}$$

$$= p(1-p) \frac{2}{(1-(1-p))^3}$$

$$= \frac{2p(1-p)}{p^3}$$

$$= \frac{2(1-p)}{p^2}$$

$$\Rightarrow \mathbb{E}[X^2] - \mathbb{E}[X] = \frac{2}{p^2} - \frac{2}{p}$$

$$\Rightarrow \quad \mathbb{E}[X^2] = \frac{2}{p^2} - \frac{2}{p} + \mathbb{E}[X]$$

$$= \frac{2}{p^2} - \frac{1}{p}$$

$$\Rightarrow \quad VAR[X] = \frac{2}{p^2} - \frac{1}{p} - \left(\frac{1}{p}\right)^2$$

$$= \frac{1}{p^2} - \frac{1}{p}$$

$$= \frac{1-p}{p^2}$$

Exercise 5 If $\mathrm{Ran}(N) = \{1, 2 \ldots\}$ show that

$$\mathbb{E}[N] = \sum_{n=1}^{\infty} P(N \geq n)$$

Start with right side:

$$\sum_{n=1}^{\infty} \mathbb{P}(N \geq n) = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} \mathbb{P}(X = k)$$

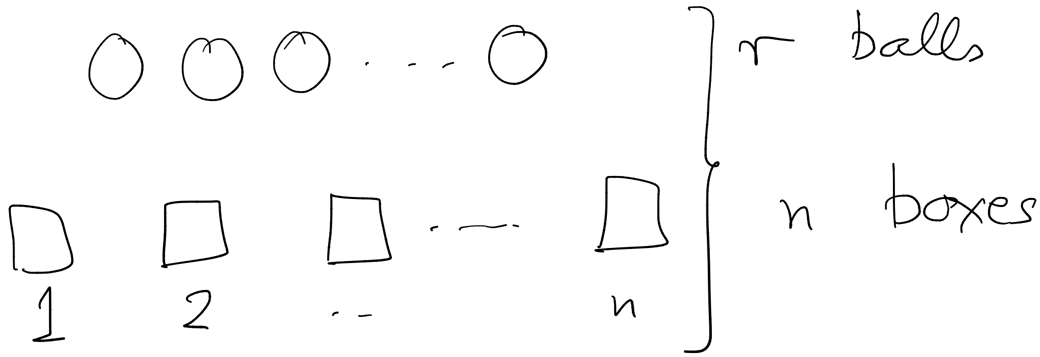
change odr of sum

$$= \sum_{k=1}^{\infty} \sum_{n=1}^{k} \mathbb{P}(X = k)$$

$$= \sum_{k=1}^{\infty} k \, \mathbb{P}(X = k)$$

$$= \mathbb{E}[X]$$

$\boxed{\textbf{Exercise 6}}$ Randomly distribute r balls in n boxes. Find the probability that the first box is empty. Find the probability that the first two boxes are both empty.



$$\mathbb{P}(\text{a given ball does not go in Box } 1) = 1 - \frac{1}{n}$$

$$\Rightarrow \mathbb{P}(\text{box } \#1 \text{ is empty})$$
$$= \mathbb{P}(\text{no balls go in box } \#1)$$
$$= \mathbb{P}(\text{first ball not in box } 1)$$
$$\cdot \mathbb{P}(\text{second ball not in box } 1)$$
$$\vdots$$
$$\cdot \mathbb{P}(r^{\text{th}} \text{ ball not in box } 1)$$

(this is a product because the events are independent)

$$= \left(1 - \frac{1}{n}\right)^r$$

$\mathbb{P}$(given ball does not go in Box 1 or Box 2)

$$= \frac{n-2}{n} = 1 - \frac{2}{n}$$

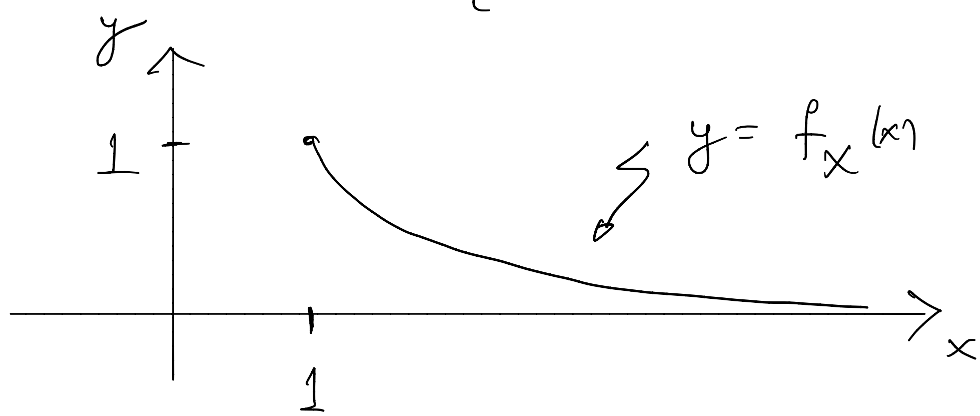
$\Rightarrow$ by same reasoning

$\mathbb{P}$(Boxes 1 and 2 are empty)

$$= \left(1 - \frac{2}{n}\right)^r$$

Exercise 7 The current in a resistor is a random variable X. The pdf of X is $f(x) = e^{-(x-1)}$ for $x \geq 1$. The power dissipated in the resistor is $Y = X^2$. Find the pdf of Y.

pdf of X: $f_X(x) = \begin{cases} e^{-(x-1)} & x \geq 1 \\ 0 & \text{else} \end{cases}$



let $Y = X^2$. What is f_Y?

First find cdf of Y:

$$F_Y(y) = \mathbb{P}(Y \leq y)$$

$$= \mathbb{P}(X^2 \leq y)$$

$$= \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y})$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\, dx$$

If $y < 1 \Rightarrow F_Y(y) = 0$

If $y \geq 1$

$$\Rightarrow F_Y(y) = \int_1^{\sqrt{y}} e^{-(x-1)}\,dx$$

$$= \int_0^{\sqrt{y}-1} e^{-u}\,du \qquad \begin{cases} u = x-1 \\ du = dx \end{cases}$$

$$= 1 - e^{-(\sqrt{y}-1)}$$

$$\Rightarrow F_Y(y) = \begin{cases} 0 & y < 1 \\ 1 - e^{-(\sqrt{y}-1)} & y \geq 1 \end{cases}$$

$$\Rightarrow f_Y(y) = \frac{d}{dy} F_Y(y)$$

$$= \begin{cases} \frac{1}{2\sqrt{y}}\, e^{-(\sqrt{y}-1)} & y \geq 1 \\ 0 & \text{else} \end{cases}$$

Exercise 8 Derive the formula

$$\mathrm{VAR}[X_1 + X_2 + \cdots + X_n] = \sum_{k=1}^{n} \mathrm{VAR}[X_k] + 2\sum_{i<j} \mathrm{COV}(X_i, X_j) \tag{1}$$

Let $\mu_i = \mathbb{E}[X_i]$, $i = 1, \ldots, n$.

$$\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mu_1 + \mu_2 + \cdots + \mu_n$$

$$\Rightarrow \mathrm{VAR}[X_1 + X_2 + \cdots + X_n]$$

$$= \mathbb{E}\left[\left(X_1 + X_2 + \cdots + X_n - (\mu_1 + \mu_2 + \cdots + \mu_n)\right)^2\right]$$

$$= \mathbb{E}\left[\left((X_1 - \mu_1) + (X_2 - \mu_2) + \cdots + (X_n - \mu_n)\right)^2\right]$$

$$= \mathbb{E}\left[(X_1 - \mu_1)^2\right] + \mathbb{E}\left[(X_2 - \mu_2)^2\right] + \cdots$$

$$+ 2\mathbb{E}\left[(X_1 - \mu_1)(X_2 - \mu_2)\right]$$

$$+ 2\mathbb{E}\left[(X_1 - \mu_1)(X_3 - \mu_3)\right] + \cdots$$

$$+ 2\mathbb{E}\left[(X_{n-1} - \mu_{n-1})(X_n - \mu_n)\right]$$

$$= VAR[X_1] + VAR[X_2] + \cdots + VAR[X_n]$$

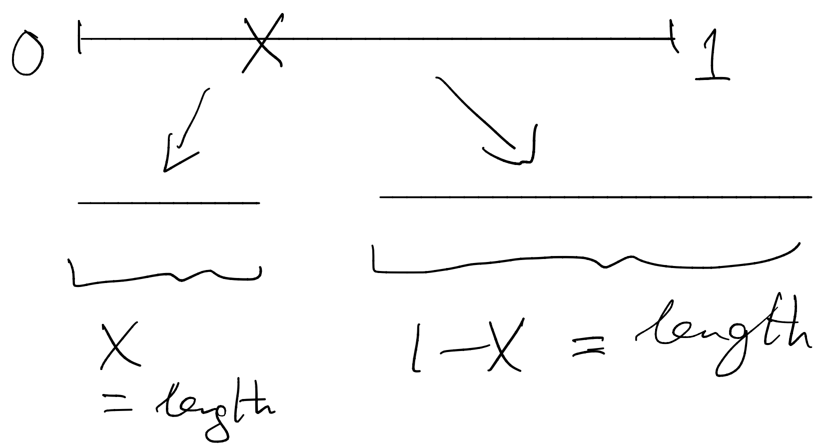
$$+ 2\,COV[X_1, X_2] + 2\,COV[X_1, X_3]$$

$$+ \cdots + 2\,COV[X_{n-1}, X_n]$$

$$= \sum_{i=1}^{n} VAR[X_i]$$

$$+ 2 \sum_{1 \le i < j \le n} COV[X_i, X_j]$$

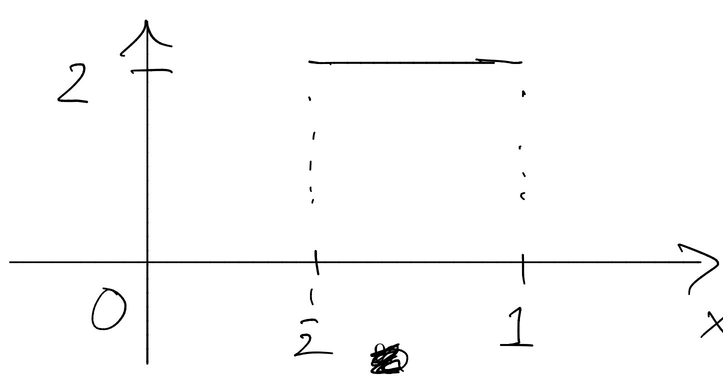
Exercise 9 We start with a stick of length 1, and break it in two pieces at a randomly chosen position (chosen uniformly over its length). Find the mean and variance of the longer piece.



L = length of larger piece

$\Rightarrow \quad L \sim U[\tfrac{1}{2}, 1]$ uniform.

$$\Rightarrow \text{pdf is} \quad f_L(x) = \begin{cases} 2 & \tfrac{1}{2} \le x \le 1 \\ 0 & \text{else} \end{cases}$$



$$\mathbb{E}[L] = \int_{\frac{1}{2}}^{1} x \cdot 2 \cdot dx = \frac{3}{4}$$

$$\mathbb{E}[L^2] = \int_{\frac{1}{2}}^{1} x^2 \cdot 2 \cdot dx = \frac{2}{3} \cdot \left(1 - \frac{1}{8}\right)$$

$$\Rightarrow VAR[L] = \mathbb{E}[L^2] - \mathbb{E}[L]^2$$

$$= \frac{7}{12} - \frac{9}{16}$$

$$= \frac{1}{48}$$

$\boxed{\textbf{Exercise 10}}$ Find a random number generator that generates uniformly on $[0, 1]$ (for example the command *rand* in Matlab). Using this generator, estimate the volume of the region under the surface

$$z = \frac{1}{3} \cosh \sqrt{x^2 + y^2}$$

and above the unit square $0 \le x \le 1$, $0 \le y \le 1$. [Note: generate three independent uniform random variables for each run, corresponding to the three coordinates of a random point in the unit cube. Do enough runs to be confident that you have an accurate estimate of the first two decimal places].

Let $U, V, W \sim U[0,1]$ independent

Let $P = (U, V, W)$ be a randomly chosen point in the unit cube $[0,1]^3$.

Then

$\mathbb{P}($ P lies below surface $)$

$= \dfrac{\text{volume of region under surface}}{\text{volume of unit cube}}$

$=$ volume of region under surface

$$= \mathbb{P}\left(W \le \frac{1}{3} \cosh \sqrt{U^2 + V^2} \right).$$

Generate N independent triplets

(U_i, V_i, W_i) $i = 1, \ldots, N$:

$$R_N = \frac{\# \left\{ i : W_i \leq \frac{1}{3} \cosh \sqrt{U_i^2 + V_i^2} \right\}}{N}$$

Then as $N \to \infty$ this ratio R_N will converge to the probability we want.

Here are results for several values of N (using Matlab):

N	Trial 1	Trial 2	Trial 3	Time per trial (sec)
1000	0.439	0.438	0.449	0.001
10000	0.453	0.45	0.457	0.003
100000	0.455	0.452	0.453	0.01
10^6	0.4531	0.4534	0.4534	0.15

Actual value of integral is

$$\int_0^1 \int_0^1 \frac{1}{3} \cosh \sqrt{x^2+y^2}\, dx\, dy = 0.4534$$

$\boxed{\textbf{Exercise 11}}$ In class we considered this problem: "An urn contains n Red balls and m Black balls. Suppose that k balls are withdrawn from the urn, and let X be the number of Red balls among these. Find $\mathbb{E}[X]$ assuming (i) replacement, and (ii) no replacement." Using the same reasoning as in class, compute $\mathrm{VAR}[X]$ assuming (i) replacement, and (ii) no replacement. [Hint: use the formula from Exercise 8 above. The answers will be different for the two cases].

$$X = R_1 + R_2 + \cdots + R_k$$

i) With replacement the $\{R_i\}$ are IID.

$$\Rightarrow \mathrm{VAR}[X] = \sum_{i=1}^{k} \mathrm{VAR}[R_i]$$

$$= k\,\mathrm{VAR}[R_1].$$

R_1	0	1
prob	$\frac{m}{n+m}$	$\frac{n}{n+m}$

$$\Rightarrow \mathrm{VAR}[R_1] = \frac{nm}{(n+m)^2}$$

$$\Rightarrow \mathrm{VAR}[X] = \frac{knm}{(n+m)^2}$$

ii) Without replacement, the $\{R_i\}$ are not independent. (Assume that $k \le n+m$).

$$VAR[X] = \sum_{i=1}^{k} VAR[R_i] + 2 \sum_{1 \le i < j \le k} COV[R_i, R_j]$$

As we showed in class, the pdf of R_i is the same for all $i = 1, \dots, k$

$$\Rightarrow VAR[R_i] = VAR[R_1] = \frac{nm}{(n+m)^2}$$

Similarly $COV[R_i, R_j]$ is the same for all pairs (i,j). Note that

$$COV[R_i, R_j] = \mathbb{E}[R_i R_j] - \mathbb{E}[R_i]\,\mathbb{E}[R_j]$$

Now $\mathbb{E}[R_i R_j] = \mathbb{P}(R_i = 1, R_j = 1)$.

As we showed in class, the probability of any sequence of colors

e.g. $R_1 B_2 B_3 R_4 R_5$

is the same if we permute it

e.g. $B_1 B_2 R_3 R_4 R_5$

So for example

$$\mathbb{P}(R_1=1, R_3=1) = \mathbb{P}(R_1=1, R_2=1, R_3=1) + \mathbb{P}(R_1=1, R_2=0, R_3=1)$$

$$= \mathbb{P}(R_1=1, R_2=1, R_3=1) + \mathbb{P}(R_1=1, R_2=1, R_3=0)$$

$$= \mathbb{P}(R_1=1, R_2=1).$$

Similarly $\mathbb{P}(R_i=1, R_j=1) = \mathbb{P}(R_1=1, R_2=1)$ for all $i<j\le k$

Now joint pdf:

R_2 \ R_1	0	1
0	$\frac{m(m-1)}{(n+m)(n+m-1)}$	$\frac{mn}{(n+m)(n+m-1)}$
1	$\frac{mn}{(n+m)(n+m-1)}$	$\frac{n(n-1)}{(n+m)(n+m-1)}$

$$\Rightarrow \mathbb{E}[R_1 R_2] = \mathbb{P}(R_1=1, R_2=1) = \frac{n(n-1)}{(n+m)(n+m-1)}$$

$$\Rightarrow \quad COV[R_1, R_2] = \frac{n(n-1)}{(n+m)(n+m-1)} - \left(\frac{n}{n+m}\right)^2$$

$$= \frac{-nm}{(n+m)^2(n+m-1)}$$

$$\Rightarrow \quad VAR[X] = k\frac{nm}{(n+m)^2} - k(k-1)\frac{nm}{(n+m)^2(n+m-1)}$$