Exam 1 Practice Questions II –solutions, 18.05, Spring 2022

Notes.

Not every possible question be covered in 11 problems. Look at the other review problems as well as the readings, psets and class problems.

Even the first 11 problems are much longer than the actual test will be,

Problem 1. *A full house in poker is a hand where three cards share one rank and two cards share another rank. How many ways are there to get a full-house? What is the probability of getting a full-house?*

Solution: We build a full-house in stages and count the number of ways to make each stage:

Stage 1. Choose the rank of the pair: $\binom{13}{1}$.

Stage 2. Choose the pair from that rank, i.e. pick 2 of 4 cards: $\binom{4}{2}$.

Stage 3. Choose the rank of the triple (from the remaining 12 ranks): $\binom{12}{1}$.

Stage 4. Choose the triple from that rank: $\binom{4}{3}$.

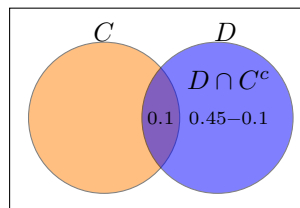Number of ways to get a full-house: $\binom{13}{1}\binom{4}{2}\binom{12}{1}\binom{4}{3}$

Number of ways to pick any 5 cards out of 52: $\binom{52}{5}$

Probability of a full house: $\dfrac{\binom{13}{1}\binom{4}{2}\binom{12}{1}\binom{4}{3}}{\binom{52}{5}} \approx 0.00144$

Problem 2. *Let C and D be two events with $P(C) = 0.25$, $P(D) = 0.45$, and $P(C \cap D) = 0.1$. What is $P(C^c \cap D)$?*

Solution: D is the disjoint union of $D \cap C$ and $D \cap C^c$.

So, $P(D \cap C) + P(D \cap C^c) = P(D)$
$\Rightarrow P(D \cap C^c) = P(D) - P(D \cap C) = 0.45 - 0.1 = \boxed{0.35.}$
(We never use $P(C) = 0.25$.)

Problem 3. *Corrupted by their power, the judges running the popular game show* America's Next Top Mathematician *have been taking bribes from many of the contestants. Each episode, a given contestant is either allowed to stay on the show or is kicked off.*

If the contestant has been bribing the judges they will be allowed to stay with probability 1. If the contestant has not been bribing the judges, they will be allowed to stay with probability 1/3.

Suppose that 1/4 of the contestants have been bribing the judges. The same contestants bribe the judges in both rounds, i.e., if a contestant bribes them in the first round, they bribe them in the second round too (and vice versa).

(a) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that they were bribing the judges?

(b) If you pick a random contestant, what is the probability that they are allowed to stay during both of the first two episodes?

(c) If you pick random contestant who was allowed to stay during the first episode, what is the probability that they get kicked off during the second episode?

Solution: The following tree shows the setting. Stay_1 means the contestant was allowed to stay during the first episode and stay_2 means the they were allowed to stay during the second.
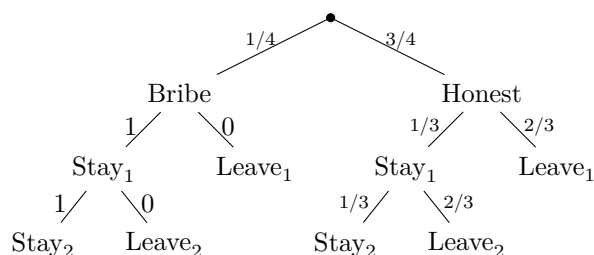
Let's name the relevant events:
B = the contestant is bribing the judges
H = the contestant is honest (not bribing the judges)
S_1 = the contestant was allowed to stay during the first episode
S_2 = the contestant was allowed to stay during the second episode
L_1 = the contestant was asked to leave during the first episode
L_2 = the contestant was asked to leave during the second episode

(a) We first compute $P(S_1)$ using the law of total probability.

$$P(S_1) = P(S_1|B)P(B) + P(S_1|H)P(H) = 1 \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{2}.$$

We therefore have (by Bayes' rule) $P(B|S_1) = P(S_1|B)\dfrac{P(B)}{P(S_1)} = 1 \cdot \dfrac{1/4}{1/2} = \dfrac{1}{2}$.

(b) Using the tree we have the total probability of S_2 is

$$P(S_2) = \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{3}$$

(c) We want to compute $P(L_2|S_1) = \dfrac{P(L_2 \cap S_1)}{P(S_1)}$.

From the calculation we did in part (a), $P(S_1) = 1/2$. For the numerator, we have (see the tree)

$$P(L_2 \cap S_1) = P(L_2 \cap S_1|B)P(B) + P(L_2 \cap S_1|H)P(H) = 0 \cdot \frac{1}{4} + \frac{2}{9} \cdot \frac{3}{4} = \frac{1}{6}$$

Therefore $\boxed{P(L_2|S_1) = \dfrac{1/6}{1/2} = \dfrac{1}{3}.}$

Problem 4. *Suppose now that events A, B and C are mutually independent with*

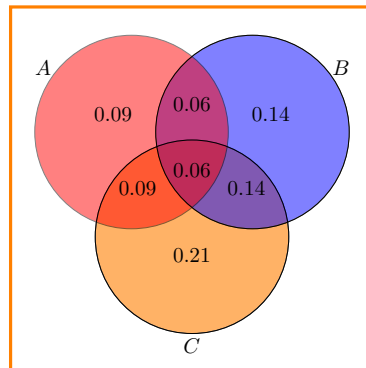$$P(A) = 0.3, \quad P(B) = 0.4, \quad P(C) = 0.5.$$

Compute the following: (Hint: Use a Venn diagram)
(i) $P(A \cap B \cap C^c)$ (ii) $P(A \cap B^c \cap C)$ (iii) $P(A^c \cap B \cap C)$

Solution: By the mutual independence we have

$$P(A \cap B \cap C) = P(A)P(B)P(C) = 0.06 \qquad P(A \cap B) = P(A)P(B) = 0.12$$
$$P(A \cap C) = P(A)P(C) = 0.15 \qquad P(B \cap C) = P(B)P(C) = 0.2$$

We show this in the following Venn diagram

Note that, for instance, $P(A \cap B)$ is split into two pieces. One of the pieces is $P(A \cap B \cap C)$ which we know and the other we compute as $P(A \cap B) - P(A \cap B \cap C) = 0.12 - 0.06 = 0.06$. The other intersections are similar.

We can read off the asked for probabilities from the diagram.
(i) $P(A \cap B \cap C^c) = 0.06$
(ii) $P(A \cap B^c \cap C) = 0.09$
(iii) $P(A^c \cap B \cap C) = 0.14$.

Problem 5. *Suppose X is a random variable with $E[X] = 5$ and $\text{Var}(X) = 2$. What is $E[X^2]$?*

Solution: Use $\text{Var}(X) = E[X^2] - E[X]^2 \Rightarrow 2 = E[X^2] - 25 \Rightarrow \boxed{E[X^2] = 27.}$

Problem 6. *(a) Suppose that X has probability density function $f_X(x) = \lambda e^{-\lambda x}$ for $x \geq 0$. Compute the cdf, $F_X(x)$.*

(b) If $Y = X^2$, compute the pdf and cdf of Y.

(a) Solution: We have cdf of X,

$$F_X(x) = \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x}.$$
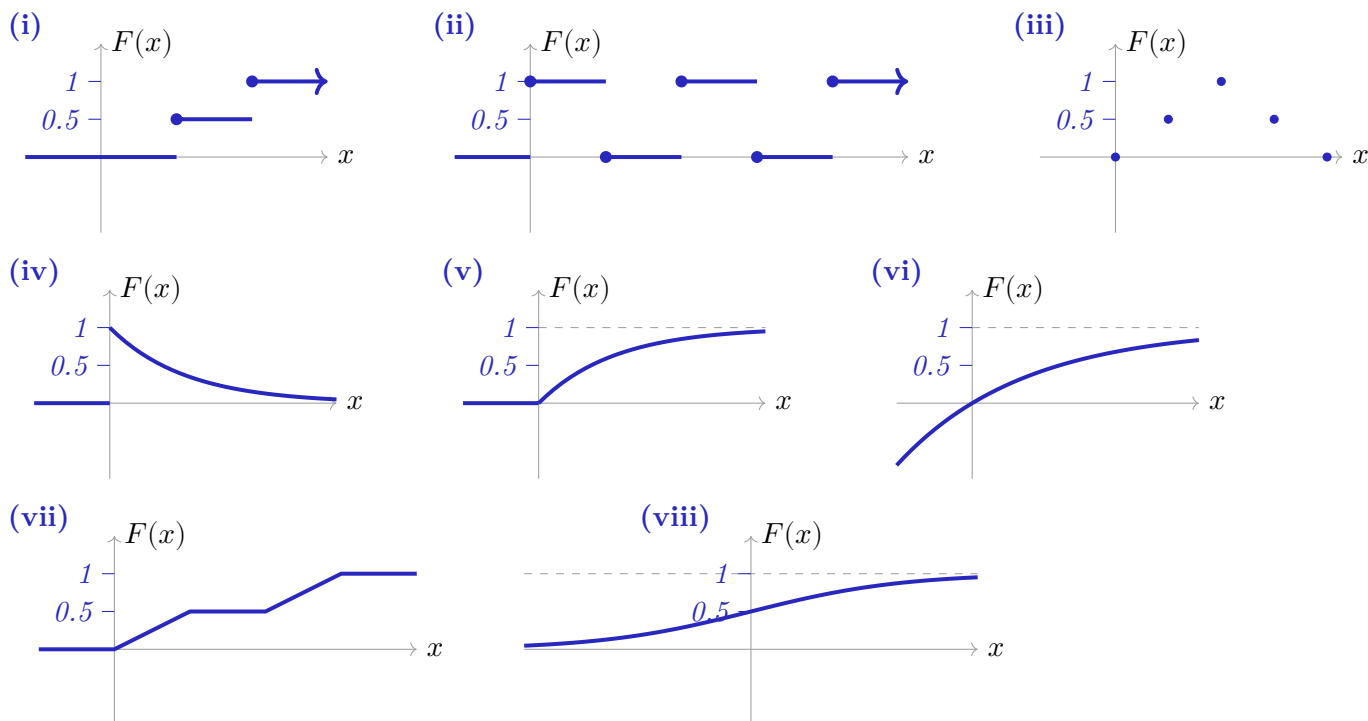
Now for $y \geq 0$, we have

(b) Solution:

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = 1 - e^{-\lambda\sqrt{y}}.$$

Differentiating $F_Y(y)$ with respect to y, we have

$$f_Y(y) = \frac{\lambda}{2} y^{-\frac{1}{2}} e^{-\lambda\sqrt{y}}.$$

Problem 7. *For each of the following say whether it can be the graph of a cdf. If it can be, say whether the variable is discrete or continuous.*

(i)

(ii)

(iii)

(iv)

(v)

(vi)

(vii)

(viii)

Solution: (i) yes, discrete, (ii) no, (iii) no, (iv) no, (v) yes, continuous (vi) no (vii) yes, continuous, (viii) yes, continuous.

Problem 8. *Compute the median for the exponential distribution with parameter λ.*

Solution: The density for this distribution is $f(x) = \lambda e^{-\lambda x}$. We know (or can compute) that the distribution function is $F(a) = 1 - e^{-\lambda a}$. The median is the value of a such that $F(a) = 0.5$. Thus, $1 - e^{-\lambda a} = 0.5 \Rightarrow 0.5 = e^{-\lambda a} \Rightarrow \log(0.5) = -\lambda a \Rightarrow \boxed{a = \log(2)/\lambda.}$

Problem 9. (Another Arithmetic Puzzle)
Let X and Y be two independent Bernoulli(0.5) random variables. Define S and T by:

$$S = X + Y \quad and \quad T = X - Y.$$

(a) Find the joint and marginal pmf's for S and T.

(b) Are S and T independent.

Solution: (a) $S = X + Y$ takes values 0, 1, 2 and $T = X - Y$ takes values -1, 0, 1.

First we make two tables: the joint probability table for X and Y and a table given the values (S, T) corresponding to values of (X, Y), e.g. $(X, Y) = (1, 1)$ corresponds to $(S, T) = (2, 0)$.

X\\Y	0	1
0	1/4	1/4
1	1/4	1/4

Joint probabilities of X and Y

X\\Y	0	1
0	0,0	1,-1
1	1,1	2,0

Values of (S, T) corresponding to X and Y

We can use the two tables above to write the joint probability table for S and T. The marginal probabilities are given in the table.

S\\T	-1	0	1	
0	0	1/4	0	1/4
1	1/4	0	1/4	1/2
2	0	1/4	0	1/4
	1/4	1/2	1/4	1

Joint and marginal probabilities of S and T

(b) No probabilities in the table are the product of the corresponding marginal probabilities. (This is easiest to see for the 0 entries.) So, S and T are not independent

Problem 10. *Let X and Y be two random variables and let r, s, t, and u be real numbers.*
(a) Show that $\text{Cov}(X + s, Y + u) = \text{Cov}(X, Y)$.

(b) Show that $\text{Cov}(rX, tY) = rt\text{Cov}(X, Y)$.

(c) Show that $\text{Cov}(rX + s, tY + u) = rt\text{Cov}(X, Y)$.

Solution: (a) First note by linearity of expectation we have $E[X + s] = E[X] + s$, thus $X + s - E[X + s] = X - E[X]$.
Likewise $Y + u - E[Y + u] = Y - E[Y]$.
Now using the definition of covariance we get

$$\text{Cov}(X + s, Y + u) = E[(X + s - E[X + s]) \cdot (Y + u - E[Y + u])]$$
$$= E[(X - E[X]) \cdot (Y - E[Y])]$$
$$= \text{Cov}(X, Y).$$

(b) This is very similar to part (a).

We know $E[rX] = rE[X]$, so $rX - E[rX] = r(X - E[X])$. Likewise $tY - E[tY] = s(Y - E[Y])$. Once again using the definition of covariance we get

$$\text{Cov}(rX, tY) = E[(rX - E[rX])(tY - E[tY])]$$
$$= E[rt(X - E[X])(Y - E[Y])]$$
$$\text{(Now we use linearity of expectation to pull out the factor of } rt)$$
$$= rtE[(X - E[X](Y - E[Y]))]$$
$$= rt\text{Cov}(X, Y)$$

(c) This is more of the same. We give the argument with far fewer algebraic details

$$\text{Cov}(rX + s, tY + u) = \text{Cov}(rX, tY) \text{ (by part (a))}$$
$$= rt\text{Cov}(X, Y) \text{ (by part (b))}$$

Problem 11. (More Central Limit Theorem)
The average IQ in a population is 100 with standard deviation 15 (by definition, IQ is normalized so this is the case). What is the probability that a randomly selected group of 100 people has an average IQ above 115?

Solution: Let X_j be the IQ of a randomly selected person. We are given $E[X_j] = 100$ and $\sigma_{X_j} = 15$.

Let $\overline{X}$ be the average of the IQ's of 100 randomly selected people. Then we know

$$E[\overline{X}] = 100 \quad \text{and} \quad \sigma_{\overline{X}} = 15/\sqrt{100} = 1.5.$$

The problem asks for $P(\overline{X} > 115)$. Standardizing we get $P(\overline{X} > 115) \approx P(Z > 10)$.
This is effectively 0.

More problems

Problem 12. *20 politicians are having a tea party, 6 Democrats and 14 Republicans. To prepare, they need to choose:*

   3 people to set the table,   2 people to boil the water,   6 people to make the scones.

Each person can only do 1 task. (Note that this doesn't add up to 20. The rest of the people don't help.)

(a) In how many different ways can they choose which people perform these tasks?

(b) Suppose that the Democrats all hate tea. If they only give tea to 10 of the 20 people, what is the probability that they only give tea to Republicans?

(c) If they only give tea to 10 of the 20 people, what is the probability that they give tea to 9 Republicans and 1 Democrat?

Solution: (a) There are $\binom{20}{3}$ ways to choose the 3 people to set the table, then $\binom{17}{2}$ ways to choose the 2 people to boil water, and $\binom{15}{6}$ ways to choose the people to make scones. So the total number of ways to choose people for these tasks is

$$\boxed{\binom{20}{3}\binom{17}{2}\binom{15}{6} = \frac{20!}{3!\,17!} \cdot \frac{17!}{2!\,15!} \cdot \frac{15!}{6!\,9!} = \frac{20!}{3!\,2!\,6!\,9!} = 775975200.}$$

(b) The number of ways to choose 10 of the 20 people is $\binom{20}{10}$ The number of ways to choose 10 people from the 14 Republicans is $\binom{14}{10}$. So the probability that you only choose 10 Republicans is

$$\frac{\binom{14}{10}}{\binom{20}{10}} = \frac{\frac{14!}{10!\,4!}}{\frac{20!}{10!\,10!}} \approx 0.00542$$

Alternatively, you could choose the 10 people in sequence and say that there is a 14/20 probability that the first person is a Republican, then a 13/19 probability that the second one is, a 12/18 probability that third one is, etc. This gives a probability of

$$\frac{14}{20} \cdot \frac{13}{19} \cdot \frac{12}{18} \cdot \frac{11}{17} \cdot \frac{10}{16} \cdot \frac{9}{15} \cdot \frac{8}{14} \cdot \frac{7}{13} \cdot \frac{6}{12} \cdot \frac{5}{11}.$$

(You can check that this is the same as the other answer given above.)

(c) You can choose 1 Democrat in $\binom{6}{1} = 6$ ways, and you can choose 9 Republicans in $\binom{14}{9}$ ways, so the probability equals

$$\frac{6 \cdot \binom{14}{9}}{\binom{20}{10}} = \frac{6 \cdot \frac{14!}{9!\,5!}}{\frac{20!}{10!\,10!}} = \frac{6 \cdot 14!\,10!\,10!}{9!\,5!\,20!}.$$

Problem 13. *More cards! Suppose you want to divide a 52 card deck into four hands with 13 cards each. What is the probability that each hand has a king?*

Solution: Let H_i be the event that the ith hand has one king. We have the conditional probabilities

$$P(H_1) = \frac{\binom{4}{1}\binom{48}{12}}{\binom{52}{13}}; \quad P(H_2|H_1) = \frac{\binom{3}{1}\binom{36}{12}}{\binom{39}{13}}; \quad P(H_3|H_1 \cap H_2) = \frac{\binom{2}{1}\binom{24}{12}}{\binom{26}{13}}$$

$$P(H_4|H_1 \cap H_2 \cap H_3) = 1$$

$$P(H_1 \cap H_2 \cap H_3 \cap H_4) = P(H_4|H_1 \cap H_2 \cap H_3)\,P(H_3|H_1 \cap H_2)\,P(H_2|H_1)\,P(H_1)$$

$$= \frac{\binom{2}{1}\binom{24}{12}\binom{3}{1}\binom{36}{12}\binom{4}{1}\binom{48}{12}}{\binom{26}{13}\binom{39}{13}\binom{52}{13}}.$$

Problem 14. *There is a screening test for prostate cancer that looks at the level of PSA (prostate-specific antigen) in the blood. There are a number of reasons besides prostate cancer that a man can have elevated PSA levels. In addition, many types of prostate cancer develop so slowly that that they are never a problem. Unfortunately there is currently no test to distinguish the different types and using the test is controversial because it is hard to quantify the accuracy rates and the harm done by false positives.*

For this problem we'll call a positive test a true positive if it catches a dangerous type of prostate cancer. We'll assume the following numbers:

Rate of prostate cancer among men over 50 = 0.0005
True positive rate for the test = 0.9
False positive rate for the test = 0.01

Let T be the event a man has a positive test and let D be the event a man has a dangerous type of the disease. Find $P(D|T)$ and $P(D|T^c)$.

Solution: You should write this out in a tree! (For example, see the solution to the next problem.)

We compute all the pieces needed to apply Bayes' rule. We're given
$P(T|D) = 0.9 \Rightarrow P(T^c|D) = 0.1, \quad P(T|D^c) = 0.01 \Rightarrow P(T^c|D^c) = 0.99.$
$P(D) = 0.0005 \Rightarrow P(D^c) = 1 - P(D) = 0.9995.$

We use the law of total probability to compute $P(T)$:

$$P(T) = P(T|D)\,P(D) + P(T|D^c)\,P(D^c) = 0.9 \cdot 0.0005 + 0.01 \cdot 0.9995 = 0.010445$$

Now we can use Bayes' rule to answer the questions:

$$P(D|T) = \frac{P(T|D)\,P(D)}{P(T)} = \frac{0.9 \times 0.0005}{0.010445} = 0.043$$

$$P(D|T^c) = \frac{P(T^c|D)\,P(D)}{P(T^c)} = \frac{0.1 \times 0.0005}{0.989555} = 5.0 \times 10^{-5}$$

Problem 15. *Let X be a discrete random variable with pmf p given by:*

x	-2	-1	0	1	2
$p(x)$	1/15	2/15	3/15	4/15	5/15

(a) Let $Y = X^2$. Find the pmf of Y.

(b) Find the value the cdf of X at -3/2, 3/4, 7/8, 1, 1.5, 5.

(c) Find the value the cdf of Y at -3/2, 3/4, 7/8, 1, 1.5, 5.

Solution: (a) Note: $Y = 1$ when $X = 1$ or $X = -1$, so

$$P(Y = 1) = P(X = 1) + P(X = -1).$$

Values y of Y	0	1	4
pmf $p_Y(y)$	3/15	6/15	6/15

(b) and **(c)** To distinguish the distribution functions we'll write F_X and F_Y.

Using the tables in part (a) and the definition $F_X(a) = P(X \le a)$ etc. we get

a	-3/2	3/4	7/8	1	1.5	5
$F_X(a)$	1/15	6/15	6/15	10/15	10/15	1
$F_Y(a)$	0	3/15	3/15	9/15	9/15	1

Problem 16. *Suppose I play a gambling game with even odds. So, I can wager b dollars and I either win or lose b dollars with probability $p = 0.5$.*

I employ the following strategy to try to guarantee that I win some money.

I bet \$1; if I lose, I double my bet to \$2, if I lose I double my bet again. I continue until I win. Eventually I'm sure to win a bet and net \$1 (run through the first few rounds and you'll see why this is the net).

If this really worked casinos would be out of business. Our goal in this problem is to understand the flaw in the strategy.

(a) *Let X be the amount of money bet on the last game (the one I win). X takes values 1, 2, 4, 8, …. Determine the probability mass function for X. That is, find $p(2^k)$, where k is in {0, 1, 2, …}.*

Solution: It is easy to see that (e.g. look at the probability tree) $P(2^k) = \dfrac{1}{2^{k+1}}$.

(b) *Compute $E[X]$.*

Solution: $E[X] = \displaystyle\sum_{k=0}^{\infty} 2^k \dfrac{1}{2^{k+1}} = \sum \dfrac{1}{2} = \infty$. Technically, $E[X]$ is undefined in this case.

(c) *Use your answer in part (b) to explain why the stategy is a bad one.*

Solution: Technically, $E[X]$ is undefined in this case. But the value of ∞ tells us what is wrong with the scheme. Since the average last bet is infinite, I need to have an infinite amount of money in reserve.

This problem and solution is often referred to as the **St. Petersburg paradox**

Problem 17. Normal Distribution: *Throughout these problems, let ϕ and Φ be the pdf and cdf, respectively, of the standard normal distribution Suppose Z is a standard normal random variable and let $X = 3Z + 1$.*

(a) *Express $P(X \leq x)$ in terms of Φ*

Solution: We have

$$F_X(x) = P(X \leq x) = P(3Z + 1 \leq x) = P\left(Z \leq \frac{x-1}{3}\right) = \Phi\left(\frac{x-1}{3}\right).$$

(b) *Differentiate the expression from (a) with respect to x to get the pdf of X, $f(x)$. Remember that $\Phi'(z) = \phi(z)$ and don't forget the chain rule*

Solution: Differentiating with respect to x, we have

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{1}{3}\phi\left(\frac{x-1}{3}\right).$$

Since $\phi(x) = (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}}$, we conclude

$$f_X(x) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2 \cdot 3^2}},$$

which is the probability density function of the $N(1, 9)$ distribution. **Note:** The arguments in (a) and (b) give a proof that $3Z+1$ is a normal random variable with mean 1 and variance 9. See Problem Set 3, Question 5.

(c) *Find $P(-1 \leq X \leq 1)$*

Solution: We have

$$P(-1 \leq X \leq 1) = P\left(-\frac{2}{3} \leq Z \leq 0\right) = \Phi(0) - \Phi\left(-\frac{2}{3}\right) \approx 0.2475$$

(d) *Recall that the probability that Z is within one standard deviation of its mean is approximately 68%. What is the probability that X is within one standard deviation of its mean?*

Solution: Since $E[X] = 1$, $\text{Var}(X) = 9$, we want $P(-2 \leq X \leq 4)$. We have

$$P(-2 \leq X \leq 4) = P(-3 \leq 3Z \leq 3) = P(-1 \leq Z \leq 1) \approx 0.68.$$

Problem 18. *Toss a fair coin 3 times. Let X = the number of heads on the first toss, Y the total number of heads on the last two tosses, and F the number of heads on the first two tosses.*

(a) Give the joint probability table for X and Y. Compute $\text{Cov}(X, Y)$.

Solution: (a) X and Y are independent, so the table is computed from the product of the known marginal probabilities. Since they are independent, $\text{Cov}(X, Y) = 0$.

$Y \backslash X$	0	1	P_Y
0	1/8	1/8	1/4
1	1/4	1/4	1/2
2	1/8	1/8	1/4
P_X	1/2	1/2	1

(b) Give the joint probability table for X and F. Compute $\text{Cov}(X, F)$.

Solution: (b) The sample space is $\Omega = $ {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}.

$P(X = 0, F = 0) = P(\{TTH, TTT\}) = 1/4$.

$P(X = 0, F = 1) = P(\{THH, THT\}) = 1/4$.

$P(X = 0, F = 2) = 0$.

$P(X = 1, F = 0) = 0$.

$P(X = 1, F = 1) = P(\{HTH, HTT\}) = 1/4$.

$P(X = 1, F = 2) = P(\{HHH, HHT\}) = 1/4$.

$F \backslash X$	0	1	P_F
0	1/4	0	1/4
1	1/4	1/4	1/2
2	0	1/4	1/4
P_X	1/2	1/2	1

$\text{Cov}(X, F) = E[XF] - E[X]E[F]$.

$E[X] = 1/2$, $E[F] = 1$, $E[XF] = \sum x_i y_j p(x_i, y_j) = 3/4$.

$\Rightarrow \text{Cov}(X, F) = 3/4 - 1/2 = \boxed{1/4.}$

MIT OpenCourseWare

https://ocw.mit.edu

18.05 Introduction to Probability and Statistics

Spring 2022