CPS 5310 Spring 2019
Homework and Lab Assignments

Homework 1 (Due Thursday, 1/31)

Refer to the *Introduction to Probability Models* book (10th Edition) by Sheldon Ross.

(a) Practice:
 Chapter 1 #1, 2, 3, 4, 11
 Chapter 2 # 1, 3, 5, 7

(b) Write up solutions for the following problems to be turned in on the due date:
 Chapter 1 #33
 Chapter 2, # 2, 4, 12, 16

Homework 2 (Due Thursday, 2/7)

(a) Read Chapter 2 of the *Statistics* book by Agresti and Franklin. In particular, make sure you get the meanings of the following terms for a set of numerical data: mean, variance, standard deviation, median, quartiles, five number summary, interquartile range, outlier. Also learn how to visualize data with dot plots, stem and leaf plots, pie charts, box plots, and histograms.

(b) Practice:
 Section 2.1 # 2.6
 Section 2.4 #2.49
 Section 2.5 #2.74
 Section 2.6 #2.89
 Chapter Problems #2.116

(c) Write up solutions for the problems below and turn in on the due date:
 Section 2.2 #2.16
 Section 2.3 #2.45
 Section 2.4 #2.48
 Section 2.5 #2.70

(d) Learn a few basic commands in R to create a data vector or matrix, to read a data file, to perform basic arithmetic operations, to get descriptive summaries and graphical visualizations like those listed in part (a). [Note: Your TA, Ms. Sumi Dey, will cover these during the lab on Tuesday, February 5. Please install R-Studio on your own laptop and bring it to the lab. Or you can use R-Studio on the lab computers.]

(e) Work out problem #2.117 using R and put your answers in a single document. The data is in the file central_park_yearly_temps.txt. Do NOT mix your answers with the R commands and outputs, which should be put in an appendix at the end of the document. Turn in a copy by email to Sumi (sdey2@miners.utep.edu) on the due date.

Homework 3 (Due Thursday, 2/14)

Refer to the *Introduction to Probability Models* book by Sheldon Ross.

(a) Practice:

Chapter 2 # 14, 22, 23, 33, 38

(b) Write up solutions for the following problems to be turned in on the due date:

Chapter 2, # 34, 40, 49, 52, 53

Homework 4 (Due Thursday, 2/21)

Refer to the *Introduction to Probability Models* book by Sheldon Ross.

(a) Practice:

Chapter 2 # 29, 36, 41, 42, 50

(b) Write up solutions for the following problems to be turned in on the due date:

Chapter 2, # 37, 45, 51, 54, 73

Solutions to Homework 3

Due to Panfeng Liang

HW 3 panfeng liang

50/50

34. (a) By properties of cumulative probability function

$$F(\infty) = \int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_0^2 C(4x-2x^2)\,dx = C\left[2x^2 - \frac{2}{3}x^3\right]\Bigg|_0^2$$

$$= C\left[2\times 4 - \frac{2}{3}\times 8\right]$$

$$= \frac{8C}{3} = 1$$

10

Thus, $C = 3/8$

(b) $$P(1/2 \le X \le 3/2) = \frac{3}{8}\int_{\frac{1}{2}}^{\frac{3}{2}} (4x-2x^2)\,dx$$

$$= \frac{3}{8}\left(2x^2 - \frac{2}{3}x^3\right)\Bigg|_{\frac{1}{2}}^{\frac{3}{2}}$$ 1.5 0.5

$$= 0.6875$$

$$= \frac{11}{16}$$

40. The distribution table can be shown as below:

# of Games	probability
4	$p^4 + (1-p)^4$
5	$\binom{4}{3} p^3(1-p)p + \binom{4}{3}(1-p)^3 p(1-p)$
6	$\binom{5}{3} p^3(1-p)^2 p + \binom{5}{3}(1-p)^3 p^2(1-p)$
7	$\binom{6}{3} p^3(1-p)^3 p + \binom{6}{3}(1-p)^3 p^3(1-p)$

10

Expected value $= 4[p^4+(1-p)^4] + 20[p^4(1-p)+(1-p)^4p]$
$+ 60[p^4(1-p)^2+(1-p)^4p^2] + 140[p^4(1-p)^3+(1-p)^4p^3]$

plug in $p=1/2$, $E = 5.8125$

49. $E(X^2)$, $[E(X)]^2$

$$E(X^2) - [E(X)]^2 = var(X) \geq 0$$

when variance = 0, for example, X is a constant, we have equality.

10

52 (a) $F_M(x) = x^n$

$f_M(x) = nx^{n-1}$

$E(M) = \int_0^1 x f_M(x)\,dx = \int_0^1 n x^n dx = \left.\frac{nx^{n+1}}{n+1}\right|_0^1$

$= \frac{n}{n+1}$

10

(b) $\int_{-1}^{1} c(1-x^2)\,dx$

$= c\left(x - \frac{x^3}{3}\right)\Big|_{-1}^{1} = c\left[\left(1-\frac{1}{3}\right) - \left(-1+\frac{1}{3}\right)\right] = 1$

$\frac{2}{3}$ $\quad -\frac{2}{3}$

$c = \frac{3}{4}$

$E(x) = \int_{-1}^{1} \frac{3}{4} x(1-x^2)\,dx = \frac{3}{4}\left(\frac{x^2}{2} - \frac{x^4}{4}\right)\Big|_{-1}^{1}$

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

(c) $E(x) = \int_0^2 cx(4x-2x^2)\,dx = c\int_0^2 (4x^2-2x^3)\,dx$

$= c\left[\frac{4x^3}{3} - \frac{x^4}{2}\right]\Big|_0^2 = {}_{\frac{3}{8}}c\left(\frac{32}{3} - 8\right)$

$c = \frac{3}{8}, \quad E(x) = 1$

53. $X \sim U(0,1)$

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

$$E(X^n) = \int_0^1 x^n \cdot 1\, dx = \left.\frac{x^{n+1}}{n+1}\right|_0^1 = \frac{1}{n+1}$$

10

$$E(X^{2n}) = \int_0^1 x^{2n}\, dx = \left.\frac{x^{2n+1}}{2n+1}\right|_0^1 = \frac{1}{2n+1}$$

$$\operatorname{var}(x^n) = E(x^{2n}) - [E(x^n)]^2$$

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

Solutions to Homework 4

Due to Yi Xie

49/50

CPS 5310 HW 4

Yi Xie

February 15, 2019

Q37

Let $X_1, X_2, \dots, X_n$ be independent random variables, each having a uniform distribution over (0,1). Let $M =$ maximum $(X_1, X_2, \dots, X_n)$. Show that the distribution function of M, $F_M(\cdot)$, is given by $F_M(x) = x^n, 0 \le x \le 1$

What is the probability density function of M?

Answer

a

Since $X_1, X_2, \dots, X_n$ be independent random variables, each having a uniform distribution over (0,1), so we have $F_{X_1}(x) = P(X_1 \le x) = F_{X_2}(x) = P(X_2 \le x) = \cdots = F_{X_n}(x) = P(X_n \le x) = x$.

$P(M \le x) = P(X_1 \le x, X_2 \le x, \cdots, X_n \le x) = P(X_1 \le x) \cdot P(X_2 \le x) \cdots P(X_n \le x) = x^n \; 0 \le x \le 1$.

The probability density function of M is $f_M(x) = \frac{d}{dx}F_M(X) = \frac{d}{dx}x^n = \begin{cases} nx^{n-1}, & 0<x<1 \\ 0 & \text{otherwise} \end{cases}$

Q45

10

A total of r keys are to be put, one at a time, in k boxes, with each key independently being put in box i with probability p_i, $\sum_{i=1}^{k} p_i = 1$. Each time a key is put in a nonempty box, we say that a collision occurs. Find the expected number of collisions.

Answer

Let N_i denote the number of keys in the i-th box and $\sum_{i=1}^{n} N_i = r$. Let I be a indicator function

$$I(N_i = 0) = \begin{cases} 1, \text{if } N_i = 0 \\ 0, \text{otherwise} \end{cases}$$

Then if $N_i \ge 1$, there are $N_i - 1 + I(N_i = 0) = N_i - 1$ times collisions in the i-th box, if $N_i = 0$, then there is $N_i - 1 + I(N_i = 0) = 0 - 1 + 1 = 0$ colllision in the i-th box.

So in the i-th box there are $N_i - 1 + I(N_i = 0)$ collisions. $X = \sum_{i=1}^{k}(N_i - 1 + I(N_i = 0))$ is the total collisions in all the box. Take expectation on both side, then we have $E(X) = E[\sum_{i=1}^{k}(N_i - 1 + I(N_i = 0))] = E[\sum_{i=1}^{k} N_i] - \sum_{i=1}^{k} E[1] + \sum_{i=1}^{k} E[I(N_i = 0)] = r - k + \sum_{i=1}^{k} E[I(N_i = 0)]$.

$E[I(N_i = 0)] = 1 \cdot P(I(N_i = 0) = 1) + 0 \cdot P(I(N_i = 0) = 0) = 1 \cdot (1 - p_i)^r$, so $E[X] = r - k + \sum_{i=1}^{k}(1 - p_i)^r$.

Q51

A coin, having probability p of landing heads, is flipped until a head appears for the r th time. Let N denote the number of flips required. Calculate $E[N]$. Hint: There is an easy way of doing this. It involves writing N as the sum of r geometric random variables.

Answer

10

The Let experiment A = flipping a coin until a head appears for the r th time. Let experiment B = flipping a coin until we get a head, then experiment A = do experiment B r times.

Let N_i denote the number of flips required to get a head $i = 1,2,3, \ldots r$. Then $N = \sum_{i=1}^{r} N_i$ and $N_1, N_2, \ldots, N_r$ are independent, each having a geometric distribution with parameter p.

So $E[N] = E[\sum_{i=1}^{r} N_i] = \sum_{i=1}^{r} E[N_i] = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}$.

Q54

Let X and Y each take on either the value 1 or -1. Let

10

$$p(1,1) = P\{X = 1, Y = 1\},$$

$$p(1,-1) = P\{X = 1, Y = -1\},$$

$$p(-1,1) = P\{X = -1, Y = 1\},$$

$$p(-1,-1) = P\{X = -1, Y = -1\}$$

Suppose that $E[X] = E[Y] = 0$. Show that (a) $p(1,1) = p(-1,-1)$; (b) $p(1,-1) = p(-1,1)$. Let $p = 2p(1,1)$. Find (c) $Var(X)$; (d) $Var(Y)$; (e) $Cov(X,Y)$.

Answer

(a) $E[X] = E[Y]$=0, so $0 = E[X] = 1 \cdot p(1{,}1) + 1 \cdot p(1,-1) - 1 \cdot p(-1{,}1) - 1 \cdot p(-1,-1) = p(1{,}1) + p(1,-1) - p(-1{,}1) - p(-1,-1)$.

$0 = E[Y] = 1 \cdot p(1{,}1) - 1 \cdot p(1,-1) + 1 \cdot p(-1{,}1) - 1 \cdot p(-1,-1)$
$= p(1{,}1) - p(1,-1) + p(-1{,}1) - p(-1,-1)$

$0 = E[X] + E[Y]$
$= p(1{,}1) + p(1,-1) - p(-1{,}1) - p(-1,-1) + p(1{,}1) - p(1,-1) + p(-1{,}1) - p(-1,-1) = 2p(1{,}1) - 2p(-1,-1)$

which implies $p(1{,}1) - p(-1,-1) = 0 \Rightarrow p(1{,}1) = p(-1,-1)$

(b) from (a) we have $p(1{,}1) = p(-1,-1)$, and $0 = E[X] = p(1{,}1) + p(1,-1) - p(-1{,}1) - p(-1,-1) = p(1,-1) - p(-1{,}1)$, so we have $p(1,-1) - p(-1{,}1) = 0 \Rightarrow p(1,-1) = p(-1{,}1)$.

(c) $Var(X) = E[X^2] - (E[X])^2 = E[X^2] = 1[p(1{,}1) + p(1,-1) + p(-1{,}1) + p(-1,-1)] = 1$ (No matter $X = 1$ or $X = -1$, $X^2 = 1$).

(d) $Var(Y) = E[Y^2] - (E[Y])^2 = E[Y^2] = 1[p(1{,}1) + p(1,-1) + p(-1{,}1) + p(-1,-1)] = 1$ (No matter $Y = 1$ or $Y = -1$, $Y^2 = 1$).

(e) $Cov(X,Y) = E[XY] - E[X] \cdot E[Y] = E[XY] = 1 \cdot p(1{,}1) - 1 \cdot p(1,-1) - 1 \cdot p(-1{,}1) + 1 \cdot p(-1,-1)$. from (a) and (b) we know that $p(1{,}1) = p(-1,-1)$ and $p(1,-1) = p(-1{,}1)$. $p = 2p(1{,}1)$, so $p(1,-1) = p(-1{,}1) = \frac{1}{2} - \frac{p}{2}$. So $Cov(X,Y) = \frac{p}{2} - (\frac{1}{2} - \frac{p}{2}) - (\frac{1}{2} - \frac{p}{2}) + \frac{p}{2} = 2p - 1$.

Q73

For the multinomial distribution (Exercise 17), let N_i denote the number of times outcome i occurs. Find

(a) $E[N_i]$;

(b) $Var(N_i)$;

(c) $Cov(N_i, N_j)$;

(d) Compute the expected number of outcomes that do not occur.

10

Answer

Let's assume that the total number of experiments is n, in each one of the experiment, the outcome i occurs with probability p_i, so we can see that N_i has a binomial distribution with parameter n and p_i.

(a) $E[N_i] = n \times p_i = np_i$

(b) $Var(N_i) = np_i(1 - p_i)$

Let $X_l = 1$ if in the $l - th$ experiment the outcome i occured, otherwise $X_l = 0$, for $l = 1,2,3, \dots, n$. $Y_m = 1$ if in the $m - th$ experiment the outcome j occured, otherwise $Y_m = 0$, for $m = 1,2,3, \dots, n$

So $N_i = \sum_{l=1}^{n} X_l$ and $N_j = \sum_{m=1}^{n} Y_m$.

(c) $Cov(N_i, N_j) = Cov(\sum_{l=1}^{n} X_l, \sum_{m=1}^{n} Y_m) = \sum_{l=1}^{n} \sum_{m=1}^{n} Cov(X_l, Y_m)$. Based on the condition, we know that if $l \neq m$ then $Cov(X_l, Y_m) = 0$, so $Cov(N_i, N_j) = \sum_{l=1}^{n} Cov(X_l, Y_l)$.

$Cov(X_l, Y_l) = E[X_l Y_l] - E[X_l]E[Y_l]$, and $X_l Y_l = 0$ (they can't happen at the same time in an experiment). $E[X_l] = p_i$, $E[Y_l] = p_j$. Now we can see that $Cov(X_l, Y_l) = E[X_l Y_l] - E[X_l]E[Y_l] = 0 - p_i p_j = -p_i p_j$, so $Cov(N_i, N_j) = \sum_{l=1}^{n} Cov(X_l, Y_l) = n * (-p_i p_j) = -np_i p_j$.

(d) Let

$$I(N_i = 0) = \begin{cases} 1, \text{if } N_i = 0 \\ 0, \text{otherwise} \end{cases}$$

So the number of outcomes that do not occur is equal to $\sum_{i=1}^{r} I(N_i = 0)$, then $E[\sum_{i=1}^{r} I(N_i = 0)] = \sum_{i=1}^{r} E[I(N_i = 0)]$, and we know that $P(N_i = 0) = (1 - p_i)^n$, so $E[I(N_i = 0)] = (1 - p_i)^n$, now we can see that the expected number of outcomes that do not occur = $\sum_{i=1}^{r} E[I(N_i = 0)] = \sum_{i=1}^{r} (1 - p_i)^n$