Exercises: Computational Probability and Statistics
CIS 2033, Section 002

April 16, 2015

1 Chapter 2, 3

Keywords: sample space, events, probability, conditional probability, multiplication rule, law of total probability, Bayes' rule

2.1 If a coin is tossed three times, what is the probability for each of the following events to happen?
(*Hint.* sample space $\Leftrightarrow$ set, events $\Leftrightarrow$ subset, probability of events $\Leftrightarrow$ area.)

  a). You get 3 heads.

  b). You get 2 heads.

  c). The first toss is head, the second toss is tail.

  d). The first two toss are heads.

3.1 Two cards are arbitrarily drawn (without replacement) from a standard 52-card deck (with 13 heart cards). Let $S1$ be the event "the first one is a heart," and $S2$ "the second one is a heart."
(*Hint.* conditional probability, multiplication rule, law of total probability.)

  a). Decide $P(S1), P(S2|S1)$, and $P(S2|S1^c)$ using the division rule.

  b). Compute $P(S2)$ using $P(S2|S1), P(S2|S1^c), P(S1)$, and $P(S1^c)$.

  c). Are $S1$ and $S2$ independent of each other? How do you know that?

3.2 There are two bags, each of which contains a mixture of red and blue balls. Let A be the event you select Bag 1 and B that you select Bag 2. Suppose $P(A) = 1/3, P(B) = 2/3$. Let C is the event that you get a red ball.

a). Describe in words what is $P(C|A), P(C^c|A), P(C), P(C^c)$.

b). Suppose you know $P(C|A) = 1/4, P(C|B) = 1/3$. Now you get a red ball, which bag does this ball most probably come from? (*Hint*. Bayes' rule.)

2   Chapter 4

Keywords: Bernoulli, Binomial, Geometric distribution, Probability Mass Function and Distribution Function

4.1   If a fair coin is tossed three times. Let X which denotes the total number of heads is a random variable.

a). What's the name of the distribution of X?

b). Write down Probility Mass Function $p(a)$ and Distribution Function $F(a)$.

4.2   Let X have a $Geo(p)$ distribution. For $n \geq 0$, show that $P(X > n) = (1 - p)^n$

$$P(X > n) = 1 - P(X \leq n)$$
$$= 1 - \sum_{k=1}^{n} P(X = k)$$
$$= 1 - \sum_{k=1}^{n} (1 - p)^{k-1} p$$
$$= 1 - p - (1 - p)p - (1 - p)^2 p - \ldots - (1 - p)^{n-1} p$$
$$= (1 - p)^2 - (1 - p)^2 p - \ldots - (1 - p)^{n-1} p$$
$$= (1 - p)^n$$

In other words, $P(X > n)$ means that the the previous n times are all failures and the probability is $(1 - p)^n$.

4.3 For a geometric distribution $Geo(p)$, show that $P(X > n + k | X > k) = P(X > n)$ for $n, k = 0, 1, 2, \ldots$.

$$
\begin{aligned}
P(X > n + k | X > k) &= \frac{P(\{X > n + k\} \cap \{X > k\})}{P(X > k)} \\
&= \frac{P(X > n + k)}{P(X > k)} \\
&= \frac{(1 - p)^{n+k}}{(1 - p)^k} \qquad (\text{ use } Exercise\ 4.2) \\
&= (1 - p)^n \\
&= P(X > n)
\end{aligned}
$$

This is known as the *memoryless property*.

The property is most easily explained in terms of "waiting times." Suppose that a random variable, X, is defined to be the time elapsed in a bank local branch from 9 am on a certain day until the arrival of the first customer: thus X is the time this local branch waits for the first customer. The "memoryless" property makes a comparison between the probability distributions of the time the local branch has to wait from 9 am onwards for his first customer, and the time that the local branch still has to wait for the first customer on those occasions when no customer has arrived by any given later time: the property of memorylessness is that these distributions of "**time from now to the next customer**" are exactly the same.

$P(X > n)$ means that the local branch has to wait for n time for the first customer.

$P(X > n + k | X > k)$ means that the local branch still has to wait for n time at any specific time point k when they still haven't met the first customer.

4.4 Homework 3, Part 2. Distinguish between a Geometric Distribution and a Binomial Distribution.

3   Chapter 5

Keywords: Uniform, Pareto, Exponential, Normal distributions, Probability Density Function and Distribution Function, Quantiles.

5.1 Homework 4, Question 1.

5.2 Let X be a continuous random variable, and its (cumulative) distribution function is given by

$$
F(x) = \begin{cases} 0, & x < 0 \\[2mm] x^3, & 0 \le x \le 1 \\[2mm] 1, & x > 1 \end{cases}
$$

a). Compute $P(1/4 < X \leq 1/2)$

b). What's the probability density function $f(x)$?

4 Chapter 6

6.1 Let U have a U (0, 1) distribution.

a). Describe the pseudo code to simulate the outcome of tossing a coin using U.

b). Describe the pseudo code to simulate $Bin(4, 0.5)$

6.2 Both the Exponential and Uniform generators are broken on Tom's computer. He gets a random number $p = 2$ from $Par(2)$ distribution.
a). How can he obtain a random number following $Unif(0, 1)$ distribution by transforming p? What's the number? (4 points)
Answer: We directly plug in p into the CDF for $Par(2)$ which is $F(x) = 1 - \frac{1}{x^2}$. We get $u = F(p) = 1 - 1/4 = 3/4$.

b). Suppose the number from a) is u, how can he obtain a random number following a $Exp(2)$ distribution by transforming u? What's the number? (4 points)
Answer: For $Exp(2)$, $F(x) = 1 - e^{-2x}$, so let $u = 3/4 = F(x) = 1 - e^{-2x}$, we get
$x = -\frac{\ln \frac{3}{4}}{2}$
Because $u' = 1 - u = 1/4$ is also drawn from a $Unif(0, 1)$ distribution, the answer
$x' = -\frac{\ln \frac{1}{4}}{2}$ is also right.

5 Chapter 7, 10

10.1 Let U and V are two discrete random variables. U takes values $1, 3, 5$ and V takes values $2, 4, 6$. Given the partially filled table of joint and marginal distributions,

	u			
v	1	3	5	P(V=v)
2		$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{6}$
4	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	
6	$\frac{1}{9}$	$\frac{1}{9}$		$\frac{1}{3}$
P(U=u)	$\frac{1}{3}$		$\frac{1}{3}$	1

a Complete the table.

b Compute $E[U], E[V], Var[U], Var[V]$.
(Hint, $E[U] = \sum_a a * P(U = a), Var[U] = E[U^2] - E[U]^2, E[U^2] = \sum_a a^2 * P(U = a)$)

c Compute $E[U + V], E[UV]$.
(Hint, $E[g(U, V)] = \sum_{a,b} g(a, b) * P(U = a, V = b)$)

c Compute $Cov(U, V)$.
(Hint, $Cov(U, V) = E[UV] - E[U]E[V]$)

d Compute $\rho(U, V)$.
(Hint, $\rho(U, V) = \frac{Cov(U,V)}{\sqrt{Var[U]Var[V]}}$)

e How is U, V correlated? Are they independent?
For a)

	u			
v	1	3	5	P(V=v)
2	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{6}$
4	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$
6	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{3}$
P(U=u)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

For b)

$$\begin{aligned}
Cov(U,V) =& E[UV] - E[U]E[V] \\
E[UV] =& 1 \times 2 \times \frac{1}{18} + 1 \times 4 \times \frac{1}{6} + 1 \times 6 \times \frac{1}{9} \\
& 3 \times 2 \times \frac{1}{18} + 3 \times 4 \times \frac{1}{6} + 3 \times 6 \times \frac{1}{9} \\
& 5 \times 2 \times \frac{1}{18} + 5 \times 4 \times \frac{1}{6} + 5 \times 6 \times \frac{1}{9} \\
=& 13 \\
E[U] =& 1 \times \frac{1}{3} + 3 \times \frac{1}{3} + 5 \times \frac{1}{3} \\
=& 3 \\
E[V] =& 2 \times \frac{1}{6} + 4 \times \frac{1}{2} + 6 \times \frac{1}{3} \\
=& \frac{13}{3} \\
Cov(U,V) =& E[UV] - E[U]E[V] \\
=& 13 - 3 \times \frac{13}{3} \\
=& 0
\end{aligned}$$

For c)

$$\begin{aligned}
\rho(U,V) =& \frac{Cov(U,V)}{\sqrt{Var(U)Var(V)}} \\
Var(U) =& E[U^2] - (E[U])^2 = \left(1^2 \times \frac{1}{3} + 3^2 \times \frac{1}{3} + 5^2 \times \frac{1}{3} \right) - 3^2 = \frac{8}{3} \\
Var(V) =& E[V^2] - (E[V])^2 = \left(2^2 \times \frac{1}{6} + 4^2 \times \frac{1}{2} + 6^2 \times \frac{1}{3} \right) - \left(\frac{13}{3} \right)^2 = \frac{17}{9}
\end{aligned}$$

Since $Cov(U,V) = 0$, then $\rho(U,V) = 0$.

6   Chapter 12

12.1   The number of errors on a hard disk is modeled as a Poisson random variable with expectation one error in every Mb, that is, in every 2^{20} bytes.

a.   What is the expected number of errors in a sector of 512 bytes ?

b.   What is the probability of at least one error in a hard disk of 1 Mb ?
For a)
We denote the random variable X as the number of errors in a sector of 512 bytes, which is known as a $Pios(\mu)$. From the question, we know that $\mu = \frac{512}{2^{20}} = 2^{-11}$.
We have to compute $E[X]$, which is $\mu = 2^{-11}$.
For b)
Let the random variable Y as the number of errors in a hard disk of 1 Mb. We know that $Y \sim Pios(\lambda)$, and $\lambda = 1$. For $Pios(\lambda)$, the pmf is $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda} = \frac{1}{k!}e^{-1}$, for $k = 0, 1, 2, \ldots$. We have to compute $P(X \geq 1) = 1 - P(X = 0) = 1- = 1 - e^{-1}$.

12.2   Suppose the rate of the arriving calls at an telephone exchange is $\lambda = \frac{1}{180}$/second, let X be the number of total calls in 5 minutes.
a). What's the distribution of X. **Answer:** X is a poison distribution. The rate for unit time one minute is $\lambda' = \lambda * 60 = 1/3.$, so $\mu = \lambda * 60 * t = 5/3$. So that $X \sim Pois(\frac{5}{3})$.

b). What's the probability of observing 0 calls in 5 minutes?
Answer: $P(X = x) = \frac{(\frac{5}{3})^x}{x!}e^{-\frac{5}{3}}$. So $P(X = 0) = e^{-\frac{5}{3}} \approx 0.19$

c). What's the probability of observing more than 3 calls?
Answer: $P(X > 3) = \sum_{x=4}^{\infty} P(X = x) = 1 - \sum_{x=0}^{x=3} P(X = x) \approx 0.09$

7   Chapter 15

15.1   Given the following data, 12, 15, 34, 24, 35, 67, 89, 13, 45, 36, 89, 46, 97, 34, 56, 78, 65, 83, 59, 74, draw 3 histograms, each of which corresponds to a specified bin width. Let the first bin start at 10. You can try 3 different bin widths: 10, 30, 50.
For the bin width 10,

Bin	Count	Height
[10, 20]	3	0.015
(20, 30]	1	0.005
(30, 40]	4	0.020
(40, 50]	2	0.010
(50, 60]	2	0.010
(60, 70]	2	0.010
(70, 80]	2	0.010
(80, 90]	3	0.015
(90, 100]	1	0.005

Draw the histogram.
For the bin width 30,

Bin	Count	Height
[10, 40]	8	0.0133
(40, 70]	6	0.0100
(70, 100]	6	0.0100

Draw the histogram.
For the bin width 50,

Bin	Count	Height
[10, 60]	12	0.012
(60, 110]	8	0.008

Draw the histogram.

15.2   Given the information about the histogram with the bin width of 30 from previous question, Compute the empirical distribution function in the point of 10, 40, 70, 100.

Bin	Height
[10, 40]	0.0133
(40, 70]	0.0100
(70, 100]	0.0100

$F(10) = 0$
$F(40) = F(10) + 0.0133 \times 30 = 0 + 0.399 \approx 0.40$
$F(70) = F(40) + 0.0100 \times 30 = 0.40 + 0.30 = 0.70$
$F(100) = F(70) + 0.010 \times 30 = 0.70 + 0.30 = 1.00$

8   Chapter 16

16.1   Given the following dataset, which contains the prime numbers less than 100.

$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97$

a.   Compute the sample median.

b.   Compute the lower and upper quartiles.
For a)
n = 25, the sample median is $x_{13} = 41$.

For b)

$$q_n(p) = x_{(k)} + \alpha \left(x_{(k+1)} - x_{(k)} \right)$$
$$k = \lfloor p(n+1) \rfloor$$
$$\alpha = p(n+1) - k$$

For the lower case, p = 0.25, n = 25
$k = \lfloor 0.25(25+1) \rfloor = 6$
$\alpha = 0.25(25+1) - 6 = 0.5$
$x_{(6)} = 13, x_{(7)} = 17$
$q_{25}(0.25) = x_{(6)} + \alpha \left(x_{(7)} - x_{(6)} \right) = 13 + 0.5(17 - 13) = 15$
For the upper case, p = 0.75, n = 25
$k = \lfloor 0.75(25+1) \rfloor = 19$
$\alpha = 0.75(25+1) - 19 = 0.5$
$x_{(19)} = 67, x_{(20)} = 71$
$q_{25}(0.75) = x_{(19)} + \alpha \left(x_{(20)} - x_{(19)} \right) = 67 + 0.5(71 - 67) = 69$

16.2 Given the following dataset, which contains Fibonacci numbers less than 100.

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89$$

a. Compute the standard deviation.

b. Compute the MAD.

For a)

n = 12. We first compute the sample mean $\bar{x}_n = \frac{0+1+1+2+3+5+8+13+21+34+55+89}{12} = \frac{232}{12} \approx 19.33$

Then, we compute the standard deviation

$$s_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2}$$

$$\sum_{i=1}^{n} (x_i - \bar{x}_n)^2$$
$$= (0 - 19.33)^2 + (1 - 19.33)^2 + (1 - 19.33)^2 + (2 - 19.33)^2 +$$
$$(3 - 19.33)^2 + (5 - 19.33)^2 + (8 - 19.33)^2 + (13 - 19.33)^2 +$$
$$(21 - 19.33)^2 + (34 - 19.33)^2 + (55 - 19.33)^2 + (89 - 19.33)^2$$
$$= 8.3307 \times 10^3$$

Then

$$s_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x}_n)^2} = \sqrt{\frac{1}{11}8.3307 \times 10^3} \approx 27.52$$

For b)

Then, to compute the *median of absolute deviation* (MAD),

$$\text{MAD}(x_1, x_2, \ldots, x_n) = \text{Med}\left(|x_1 - \text{Med}_n|, \ldots, |x_n - \text{Med}_n|\right)$$

n = 12, the sample median is $\text{Med}_n = \frac{x_6 + x_7}{2} = \frac{5+8}{2} = 6.5$

Actually, we have to compute the following dataset

$$|0-6.5|, |1-6.5|, |1-6.5|, |2-6.5|, |3-6.5|, |5-6.5|, |8-6.5|, |13-6.5|, |21-6.5|, |34-6.5|, |55-6.5|, |89-6.5|$$

$$6.5, 5.5, 5.5, 4.5, 3.5, 1.5, 2.5, 7.5, 15.5, 18.5, 49.5, 83.5$$

Now we order those datapoints

$$1.5, 2.5, 3.5, 4.5, 5.5, 5.5, 6.5, 7.5, 15.5, 18.5, 49.5, 83.5$$

Now, the median is $\text{MAD} = \frac{5.5 + 6.5}{2} = 6$.