Problem Section 5

Mon Oct 30, 2023

Learning Outcomes

The problems are designed to build conceptual understanding and problem-solving skills. The emphasis is on learning to find, evaluate and build confidence. The specific tasks include:

- Calculate expected value and variance using the PMF
- Use Chebychev's inequality to make probability statements
- Use the geometric series to calculate probabilities
- Back up and support work with relevant explanations

Exercises

1. (Mean and variance)

a. Suppose a random variable X has PMF as shown below. Find $E[X]$. Also calculate $Var[X]$.

x	-\$5	\$170
$f(x)$	37/38	1/38

We may solve for $E(X)$ and $Var(X)$ using the definitions of these quantities:

$$E(X) = \sum_x x f(x) = -5 \times \frac{37}{38} + 170 \times \frac{1}{38}.$$

Since $Var(X) = E(X^2) - E(X)^2$. We know $E(X)$, so we can now find $E(X^2)$

$$E(X^2) = \sum_x x^2 f(x) = -5^2 \times \frac{37}{38} + 170^2 \times \frac{1}{38}$$

These are calculated below.

```
E_X <- -5*37/38 + 170/38
Var_X <- (25*37/38 + 170^2/38 - E_X^2)
SD_X <- sqrt(Var_X)    #SD = sqrt(Var(X))
```

Thus we have that $E(X)$ = -\$0.39 and $SD(X)$ = \$28.01.

The SD tells us the typical deviation in earnings from the average earnings ($E(X)$). It is really high in this example. It tells us that typically our net gain will be in the interval $[E(X) - SD(X), E(X) - SD(X)] = [-28.41, 27.62]$

b. Suppose $X \sim Binom(n = 10, \pi = \frac{2}{3})$. What is the expected value of $Y = 3X - 4$?

By Theorem 7.1, we know that $E[X] = 10 \times \frac{2}{3}$. Using the linearity of expectation, we can say

$$E[Y] = 3 \times E[X] - 4 = 16.$$

c. If X denotes a temperature of a randomly selected day recorded in degrees Fahrenheit, then $Y = \frac{5}{9}X - \frac{160}{9}$ is the corresponding temperature in degrees Celsius. If the standard deviation for X is $15.7°F$, what is the standard deviation of Y?

By Lemma 7.4, we have the result that

$$Var(Y) = \left(\frac{5}{9}\right)^2 Var(X).$$

Since $Var(X) = 15.7^2$ we can say that

$$Var(Y) = \frac{25 \times 15.7 \times 15.7}{81}$$

and

$$SD(Y) = \sqrt{Var(Y)} = \frac{5 \times 15.7}{9}.$$

2. (Chebychev) Suppose X is a random variable with mean and variance both equal to 20. What can be said about $P(0 < X < 40)$?

Hint: Chebychev's inequality says that

$$P(|X - 20| \geq k\sqrt{20}) \leq \frac{1}{k^2}.$$

What would you choose for k here so you can say something about $P(0 < X < 40)$?

We have that:

$$\begin{aligned}
P(0 < X < 40) &= P(-20 < X - 20 < 20) \\
&= P(\frac{-20}{\sqrt{20}} < \frac{X - 20}{\sqrt{20}} < \frac{20}{\sqrt{20}}) \\
&= P(-\sqrt{20} < \frac{X - 20}{\sqrt{20}} < \sqrt{20}) \\
&= P(\frac{|X - 20|}{\sqrt{20}} < \sqrt{20}) \\
&= 1 - P(\frac{|X - 20|}{\sqrt{20}} \geq \sqrt{20}) \\
&= 1 - P(|X - 20| \geq \underbrace{\sqrt{20}}_{\sigma} \times \underbrace{\sqrt{20}}_{k}) \\
&\geq 1 - \frac{1}{\underbrace{\sqrt{20}^2}_{k^2}} \\
&= 0.95
\end{aligned}$$

3. (Suppose we wish to generate $X \sim Binom(n = 10, \pi = \frac{2}{3})$ subject to the constraint $X \leq 3$. Say we use the following naive algorithm to accomplish this task:

- Generate an x from a $Binom(10, \frac{2}{3})$
- Accept the value x if $x <= 3$. Otherwise reject it.

a. Calculate the acceptance probability. That is, what is the probability we will accept a value x that is generated?

   We accept x if $x \leq 3$. Thus the acceptance probability is $P(X \leq 3)$. We can find this using `pbinom`

```
pi <- pbinom(q = 3,size = 10, prob = 2/3)
```

   The acceptance probability is 0.0197.

   Let's put our calculation to the test by actually generating 1,000 binomial random variables and seeing how many we would accept.

```
set.seed(1414)

x <- rbinom(n=1000,size=10,prob=2/3)

sum(x <= 3)/1000          #fraction of accepted values
```

```
## [1] 0.023
```

b. Define a new random variable Y as the number of times we have to generate a binomial variable before we find an acceptable one. For example, if on our first try, we get $x = 2$, then $y = 0$. What is the distribution of Y? Be sure to state the distribution with the parameter specified.

   The random variable Y has a geometric distribution with success probability $\pi = 0.0197$.

c. How many x should you expect to reject? That is, what is $E[Y]$? Write the R function for calculating the probability that Y is larger than expected.

   Since the mean of a $Geom(\pi)$ random variable is $\frac{1-\pi}{\pi}$, we have $E[Y] = 49.8605$. We know from example 8.2 that if $Y \sim Geom(\pi)$ then for any integer k $P(Y \geq k) = (1-\pi)^k$. This is calculated below using this formula and also using the R function `pgeom`.

```
E_Y <- (1-pi)/pi      #E_Y = 49.8605

Prob_more_than_expected <- (1 - pi)^(ceiling(E_Y))    #P(Y >=50)

print(Prob_more_than_expected)
```

```
## [1] 0.37051
```

```
pgeom(q = E_Y, prob =pi, lower.tail = F ) #P(Y > E_Y)
```

```
## [1] 0.37051
```