

Review of Introduction to Probability and Statistics

Chris Mack, <http://www.lithoguru.com/scientist/statistics/review.html>

Homework #2 Solutions

1. Consider an untested batch of memory chips that have a known failure rate of 8% (yield = 92%).
 - a. What is the probability that exactly 4 out of a lot of 12 will fail?
 - b. What is the probability that at most 4 out of a lot of 12 will fail?
 - c. What assumptions have you made in answering the two above questions?

Solution:

- (a) Use the binomial distribution with $p = 0.08$, $n = 12$, $k = 4$

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{12}{4} (0.08)^4 (0.92)^8 = 0.0104$$

- (b) The probability that at most 4 will fail is $p_X(0) + p_X(1) + p_X(2) + p_X(3) + p_X(4)$

$$= 0.3677 + 0.3837 + 0.1835 + 0.0532 + 0.0104 = 0.9984$$

(c) The main assumption is that of independence and identical distribution: Each individual part has the same probability of failing independent of whether others in the batch are defective. This is equivalent to saying there is no “bunching” of failures – a group of failed chips occurring together and all caused by the same problem.

2. Consider a discrete random variable X with the following PMF:

$$p_X(x) = \begin{cases} \frac{1}{9}, & \text{if } x \text{ is an integer in the range of } [-4, 4] \\ 0, & \text{otherwise} \end{cases}$$

Determine the PMF for $Y = |X|$.

Solution:

A function of a random variable is a random variable with PMF

$$p_Y(y) = \sum_{x|g(x)=y} p_X(x)$$

Since $y = 1$ when $x = \pm 1$, $y = 2$ when $x = \pm 2$, etc., we will have

$$p_Y(y) = \begin{cases} 1/9 & \text{for } y = 0 \\ 2/9 & \text{for } y = 1, 2, 3, 4 \\ 0 & \text{otherwise} \end{cases}$$

3. You are at a party with 500 guests. What is the probability that exactly one other guest has the same birthday as you? For simplicity, exclude the possibility of a February 29 birthday. After calculating the result exactly, approximate the answer by using a Poisson distribution.

Solution:

Use the binomial distribution with $n = 499$, $k = 1$, and $p = 1/365$ (the that any random person has the same birthday as you) Thus,

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{499}{1} \left(\frac{1}{365}\right)^1 \left(\frac{364}{365}\right)^{499} = 0.3487$$

(I used the BINOM.DIST function in Excel to calculate this result)

For large n , the binomial distribution is well approximated by a Poisson distribution with parameter $\lambda = np$. Since $np = 499/365 = 1.367$, the Poisson approximation is

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda} = \frac{1.367^1}{1!} e^{-1.367} = 0.3484$$

Note that this number differs only in the fourth significant digit.

4. After renting a large house on vacation, the realtor gave you five unmarked and similar-looking keys for the five doors of the house. Find the PMF of the number of trials required before successfully opening the front door under these different assumptions:
- After an unsuccessful trial, you set the bad key aside so you don't try it again.
 - After an unsuccessful trial, you put the bad key back into the pool of keys so that each key has an equal likelihood of being picked.

Solution:

Let random variable X be the number of trials you need to open the door, and let K_i be the event that the i^{th} key selected opens the door.

(a) In this case, we have $p_X(1) = \mathbb{P}(K_1) = 1/5$. Getting the key right on the second try means that we didn't get the key right on the first try. Thus, $p_X(2) = \mathbb{P}(K_1^c) \mathbb{P}(K_2|K_1^c) = (4/5)(1/4) = 1/5$.

Proceeding similarly, we see that the PMF of X is $p_X(x) = 1/5$, $x = 1, 2, 3, 4, 5$.

We can also view the problem as ordering the keys in advance and then trying them in succession, in which case the probability of any of the five keys being correct is $1/5$.

(b) In this case, X is a geometric random variable with $p = 1/5$, and its PMF is

$$p_X(k) = (1 - p)^{k-1}p = \left(\frac{4}{5}\right)^{k-1} \left(\frac{1}{5}\right)$$

5. Let X be a discrete random variable with PMF

$$p_X(x) = \begin{cases} x^2/a, & \text{if } x = -3, -2, -1, 0, 1, 2, 3 \\ 0, & \text{otherwise} \end{cases}$$

- Find a and $E[X]$.
- What is the PMF of the random variable $Z = (X - E[X])^2$?
- Using the result from part (b), find the variance of X .
- Find the variance of X using the formula $\text{var}[X] = \sum_{\text{all } x} (x - E[X])^2 p_X(x)$.

Solution:

(a) Normalization requires that $\sum p_X(x) = 1$. Thus,

$$1 = \frac{1}{a} \sum_{x=-3}^3 x^2 \rightarrow a = 28.$$

We can calculate the expectation in the normal way ($E[X] = \sum x p_X(x)$) or we can recognize that $E[X] = 0$ since the PMF is symmetric about zero.

(b) For $x = z = 0$, $p_Z(0) = 0$. For all other values ($z = 1, 4$, and 9), we have

$$p_Z(z) = p_X(\sqrt{z}) + p_X(-\sqrt{z}) = \frac{z}{28} + \frac{z}{28} = \frac{z}{14}$$

(c) The variance of X will be the expectation of z .

$$E[Z] = \sum z p_Z(z) = \frac{1}{14} + \frac{4^2}{14} + \frac{9^2}{14} = 7$$

(d) Using the definition of variance,

$$\text{var}[X] = \sum_{\text{all } x} (x - E[X])^2 p_X(x) = \frac{1}{28} \sum_{x=-3}^3 x^4 = 7.$$

6. A city's temperature is modeled as a random variable with mean and standard deviation both equal to 10 degrees Celsius. A day is described as "typical" if the temperature during that day ranges within one standard deviation from the mean. What would be the temperature range for a typical day if temperature were expressed in degrees Fahrenheit?

Solution:

If X is the temperature in Celsius, the temperature in Fahrenheit is $Y = 32 + 9X/5$. Therefore,

$$E[Y] = 32 + 9E[X]/5 = 32 + 18 = 50.$$

Also $\text{var}(Y) = (9/5)^2 \text{var}(X)$ where $\text{var}(X)$, the square of the given standard deviation of X , is equal to 100. Thus, the standard deviation of Y is $(9/5)10 = 18$. Hence a normal day in Fahrenheit is one for which the temperature is in the range $[32, 68]$.

7. The St. Petersburg paradox: You toss independently a fair coin and you count the number of tosses until the first tail appears. If this number is n , you receive 2^n dollars. What is the expected amount that you will receive? How much would you be willing to pay to play this game?

Solution:

This is a geometric distribution with $p = 0.5$ and a PMF of $p_X(k) = (1 - p)^{k-1}p = \left(\frac{1}{2}\right)^k$. The payout is $Y = 2^k$ and this payout has the same PMF as X . Thus,

$$E[Y] = \sum_{k=1}^{\infty} y p_Y(k) = \sum_{k=1}^{\infty} 2^k \left(\frac{1}{2}\right)^k = \sum_{k=1}^{\infty} 1 = \infty.$$

How much would you be willing to pay to play a game with an expected payout of infinity?

8. Alice passes through four traffic lights on her way to work, and each light is equally likely to be green or red, independent of the others.
- What is the PMF, the mean, and the variance of the number of red lights that Alice encounters?
 - Suppose that each red light delays Alice by exactly two minutes. What is the variance of Alice's commuting time?

Solution:

(a) Let X be the number of red lights that Alice encounters. The PMF of X is binomial with $n = 4$ and $p = 1/2$. The mean and the variance of X are $E[X] = np = 2$ and $\text{var}(X) = np(1 - p) = 1$.

(b) Let Y = the time that Alice is delayed by the red lights. $Y = 2X$ since each red light delays Alice by 2 minutes. Since $\text{var}(Y) = 4\text{var}(X)$, and $\text{var}(X) = 1$ from part (a), $\text{var}(Y) = 4$.

9. Given the PDF for a uniformly distributed continuous random variable X shown below, derive the expectation and the variance of X .

$$f_X(x) = \begin{cases} 1/(b-a), & \text{if } a < x < b \\ 0, & \text{otherwise} \end{cases}$$

Solution:

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{b+a}{2} \\ \text{var}[X] &= \int_{-\infty}^{\infty} (x - E[X])^2 f_X(x) dx = \frac{1}{b-a} \int_a^b \left(x - \frac{b+a}{2}\right)^2 dx \\ &= \frac{1}{b-a} \int_a^b \left(x^2 - (b+a)x + \left(\frac{b+a}{2}\right)^2\right) dx = \left(\frac{1}{b-a}\right) \left(\frac{x^3}{3} - (b+a)\frac{x^2}{2} + \left(\frac{b+a}{2}\right)^2 x\right) \Big|_a^b \\ &= \left(\frac{1}{b-a}\right) \left(\frac{b^3 - a^3}{3} - (b+a)\frac{(b^2 - a^2)}{2} + \left(\frac{b+a}{2}\right)^2\right) \\ &= \left(\frac{1}{b-a}\right) \left(\frac{(b-a)(b^2 + ab + a^2)}{3} - (b+a)\frac{(b+a)(b-a)}{2} + \left(\frac{b+a}{2}\right)^2\right) \\ &= \frac{(b^2 + ab + a^2)}{3} - \frac{(b+a)^2}{4} = \frac{4(b^2 + ab + a^2)}{12} - \frac{3(b^2 + 2ab + a^2)}{12} = \frac{(b-a)^2}{12} \end{aligned}$$

An easier approach, from an algebraic standpoint, is to use $\text{var}[X] = E[X^2] - (E[X])^2$.

10. What is the CDF (cumulative distribution function) for the PDF of problem 9?

Solution:

$$F_X(x) = \int_{-\infty}^x f_X(x) dx = \frac{1}{b-a} \int_a^x dx = \frac{x-a}{b-a} \quad \text{if } a < x < b$$

For $x < a$, $F_X(x) = 0$. For $x > b$, $F_X(x) = 1$.

11. Consider the exponential PDF with (positive) parameter λ :

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

- Show that this is a legitimate PDF.
- Find the expectation and variance of X .

Solution:

(a) To show that the PDF is legitimate, we must show that it is never negative and normalizes to 1. By inspection, the PDF is never negative.

$$\int_{-\infty}^{\infty} f_X(x) dx = \int_0^{\infty} \lambda e^{-\lambda x} dx = -e^{-\lambda x} \Big|_0^{\infty} = 1$$

(b)

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^{\infty} x \lambda e^{-\lambda x} dx = -e^{-\lambda x} (x + 1/\lambda) \Big|_0^{\infty} = \frac{1}{\lambda}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = -e^{-\lambda x} \left(x^2 + \frac{2x}{\lambda} + \frac{2}{\lambda^2} \right) \Big|_0^{\infty} = \frac{2}{\lambda^2}$$

$$\text{var}[X] = E[X^2] - (E[X])^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda} \right)^2 = \frac{1}{\lambda^2}$$

12. Let X and Y be normal random variables with means 0 and 1, respectively, and variances 1 and 4, respectively.

- Find $\mathbb{P}(X < 1.5)$ and $\mathbb{P}(X < -1)$.
- Find the PDF of $(Y - 1)/2$.
- Find $\mathbb{P}(-1 < Y < 1)$.

Solution:

(a) X is a standard normal, so by using the normal table, we have $\mathbb{P}(X < 1.5) = 0.9332$. Also $\mathbb{P}(X < -1) = \mathbb{P}(X > 1) = 1 - \mathbb{P}(X < 1) = 1 - 0.8413 = 0.1587$.

(b) The random variable $(Y - 1)/2$ is obtained by subtracting from Y its mean (which is 1) and dividing by the standard deviation (which is 2), so the PDF of $(Y - 1)/2$ is the standard normal.

(c) We have, using the normal table,

$$\mathbb{P}(-1 < Y < 1) = \mathbb{P}(-1 < (Y - 1)/2 < 0) = \mathbb{P}(-1 < Z < 0) = \mathbb{P}(0 < Z < 1) = \mathbb{P}(Z < 1) - \mathbb{P}(Z < 0)$$

$$= 0.8413 - 0.5 = 0.3413.$$

13. A city's temperature is modeled as a normal random variable with mean and standard deviation both equal to 10 degrees Celsius. What is the probability that the temperature at a randomly chosen time will be less than or equal to 59 degrees Fahrenheit?

Solution:

Let X and Y be the temperature in Celsius and Fahrenheit, respectively, which are related by $X = 5(Y - 32)/9$. Therefore, 59 degrees Fahrenheit correspond to 15 degrees Celsius. So, if Z is a standard normal random variable, we have using $E[X] = 10$ and $\sigma_X = 10$,

$$P(Y \leq 59) = P(X \leq 15) = P\left(Z \leq \frac{15 - E[X]}{\sigma_X}\right) = P(Z \leq 0.5) = 0.6915.$$