

CS 70, Summer 2014 — Homework 7

Harsimran (Sammy) Sidhu, SID 23796591

August 15, 2014

Collaborators: Chonyi Lama, Jenny Pushkarskaya

Sources:

Problem 1

Bond has three choices each with a probability of $1/3$. If he goes out the door, it takes him no time to escape. The route with the AC duct will take him 2 hours and he will end up where he started. If he goes into the sewer pipe it will take 5 hours and he will end up in the same place again. So the expected value is the expected value of each times the probability of each happening.

$$E(X) = \frac{1}{3}(0) + \frac{1}{3}(2 + E(x)) + \frac{1}{3}(5 + E(x))$$

$$E(X) = \frac{2}{3} + \frac{1}{3}E(x) + \frac{5}{3} + \frac{1}{3}E(x)$$

$$E(X) = \frac{7}{3} + \frac{2}{3}E(x)$$

$$\frac{1}{3}E(X) = \frac{7}{3}$$

$$E(X) = 7$$

Problem 2

(a) $P[X = i] = \binom{n}{i} p^i (1-p)^{n-i}$

$$P[X = 3] = \binom{20}{3} (1/6)^3 (5/6)^{20-3}$$

$$P[X = 3] = \binom{20}{3} (1/6)^3 (5/6)^{17}$$

(b) $P[X \leq k] = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i}$

$$P[X \leq 3] = \sum_{i=0}^3 \binom{20}{i} (1/6)^i (5/6)^{20-i}$$

$$P[X \leq 3] = \sum_{i=0}^3 \binom{20}{i} (1/6)^i (5/6)^{20-i}$$

$$P[X \leq 3] = \binom{20}{0} (1/6)^0 (5/6)^{20} + \binom{20}{1} (1/6)^1 (5/6)^{20-1} + \binom{20}{2} (1/6)^2 (5/6)^{20-2} + \binom{20}{3} (1/6)^3 (5/6)^{20-3}$$

$$P[X \leq 3] = \left(\frac{5}{6}\right)^{20} + 20\left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^{19} + \binom{20}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{18} + \binom{20}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{17}$$

(c) The probability we don't roll a 6 on the first try is $\frac{1}{6}$ and ditto for the second and third. so the probability we roll at least 4 dice.

$$P[X > 3] = (1-p)(1-p)(1-p)$$

$$P[X > 3] = (5/6)^3$$

Problem 3

$$(a) \ P[X = i] = \frac{\lambda^i}{i!} e^{-\lambda}$$

$$P[X = 7] = \frac{(20)^7}{7!} e^{-20}$$

$$(b) \ P[X = i] = \frac{\lambda^i}{i!} e^{-\lambda}$$

$$P[X < 2] = \frac{(2)^1}{1!} e^{-2} + \frac{(2)^0}{0!} e^{-2} = \frac{2}{e^2} + \frac{1}{e^2} = \frac{3}{e^2}$$

$$(c) \ \lambda = 2 \times 5.7 = 11.4$$

$$P[X < 3] = \frac{(11.4)^2}{2!} e^{-11.4} + \frac{(11.4)^1}{1!} e^{-11.4} + \frac{(11.4)^0}{0!} e^{-11.4}$$

$$P[X < 3] = \frac{(11.4)^2}{2!} e^{-11.4} + 11.4 e^{-11.4} + e^{-11.4}$$

$$P[X > 2] = 1 - P[X < 3] = 1 - \left(\frac{(11.4)^2}{2!} e^{-11.4} + 11.4 e^{-11.4} + e^{-11.4} \right)$$

Problem 4

- (a) Let's say we have a building with n floors with m people in the elevator. The probability of a person not getting off the first floor is $(\frac{n-1}{n})$ and the probability that all m people don't get off at the first floor is $(\frac{n-1}{n})^m$. Using linearity of expectation we can say that everyone not getting off an i^{th} floor is still $(\frac{n-1}{n})^m$.

$$p = (\frac{n-1}{n})^m$$

The expected value of the floors we don't stop at is np which is $n(\frac{n-1}{n})^m$

$$E(X) = n(\frac{n-1}{n})^m$$

$$var(X) = E(X^2) - E(X)^2$$

$$var(X) = \sum_{i=1}^n E(X_i^2) + \sum_{i \neq j}^n E(X_i X_j) - E(X)^2$$

Due to the fact that X_i is always 1 or 0 we can simplify X_i^2 to X_i .

$$var(X) = \sum_{i=1}^n E(X_i) + \sum_{i \neq j}^n E(X_i X_j) - E(X)^2$$

$$var(X) = \sum_{i=1}^n E(X_i) + \sum_{i \neq j}^n p_1 p_2 - E(X)^2$$

$$var(X) = \sum_{i=1}^n E(X_i) + \sum_{i \neq j}^n p_1 p_2 - E(X)^2$$

In this case p_1 is the probability that we don't pick some floor. p_2 is the probability that we don't pick the floor from p_1 and some other floor.

Therefore p_1 is simply $(\frac{n-1}{n})^m$ and p_2 is $(\frac{n-2}{n-1})^m$

$$var(X) = np_1 + n(n-1)p_1 p_2 - (np_1)^2$$

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

- (b) If the three friends have their own sequence of books the probability that the first book is the same for all 3 is simply $\frac{1}{n^2}$ due to the fact that we don't care what book it is except that it is the same for all three. The second book would also have the same probability and so on and so forth.

$$\text{var}(X) = E(X^2) - E(X)^2$$

$$\text{var}(X) = \sum_{i=1}^n E(X_i^2) + \sum_{i \neq j}^n E(X_i X_j) - E(X)^2$$

Due to the fact that X_i is always 1 or 0 we can simplify X_i^2 to X_i .

$$\text{var}(X) = \sum_{i=1}^n E(X_i) + \sum_{i \neq j}^n E(X_i X_j) - E(X)^2$$

$$\text{var}(X) = \sum_{i=1}^n E(X_i) + \sum_{i \neq j}^n p_1 p_2 - E(X)^2$$

p_1 is the probability that the same book is chosen for that index and p_2 is the probability that the same book was chosen again but not the same book as p_1 .

$$p_1 = \frac{1}{n^2}$$

$$p_2 = \frac{1}{(n-1)^2}$$

$$\text{var}(X) = n \frac{1}{n^2} + n(n-1) \frac{1}{n^2} \frac{1}{(n-1)^2} - \left(\frac{1}{n}\right)^2$$

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

Problem 5

From the notes we know that $Var(X) = E(X^2) - E(X)^2 = \sigma^2$

Since the variance is the standard deviation squared so it must be a positive value. Therefore $E(X^2)$ must be equal in which the variance is 0 or $E(X^2)$ must be larger which would cause the variance would be larger than 0 which it only can be.

$$E(X^2) \geq E(X)^2$$

If a random variable is constant then it can only have 1 value. Therefore the variability must be 0. Using the equation we see that $E(X^2) - E(X)^2 = 0$ which can be arranged to $E(X^2) = E(X)^2$

Problem 6

- (a) There is a total 7 questions and the mean for each question is 5 points. Therefore the mean for the total will be $7 \times 5 = 35$ points.

(b) $P[X \geq \alpha] = \frac{E(X)}{\alpha}$

$$P[X \geq (60)] = \frac{E(X)}{60} = \frac{35}{60} = \frac{7}{12}$$

(c) $\text{Var}(X) = \text{var}(3x) + \text{var}(4x) = 3^2\text{var}(x) + 4^2\text{var}(x) = 3^2(1) + 4^2(1) = 25$

(d) $P[|X - \mu| \geq \alpha] \leq \frac{\text{Var}(X)}{\alpha^2}$

$$P[|X - 35| \geq (60 - 35)] \leq \frac{25}{(60 - 35)^2}$$

$$P[|X - 35| \geq 25] \leq \frac{25}{25^2} = \frac{1}{25}$$

Problem 7

$$(a) \quad n \geq \frac{\sigma^2}{\epsilon^2 \delta}$$

$$n\delta \geq \frac{\sigma^2}{\epsilon^2}$$

$$\delta \geq \frac{\sigma^2}{n\epsilon^2}$$

$$\delta \geq \frac{10}{1000 \times (.5)^2}$$

$$\delta \geq \frac{1}{25} = .04$$

So we are 96% confident.

$$(b) \quad n \geq \frac{\sigma^2}{\epsilon^2 \delta}$$

$$n \geq \frac{10}{2^2 \times (0.02)}$$

$$n \geq \frac{250}{2} = 125$$

$$(c) \quad n \geq \frac{\sigma^2}{\epsilon^2 \delta}$$

$$\epsilon^2 \geq \frac{\sigma^2}{n\delta}$$

$$\epsilon \geq \sqrt{\frac{\sigma^2}{n\delta}}$$

$$\epsilon \geq \sqrt{\frac{10}{2500 \times .1}}$$

$$\epsilon \geq \sqrt{\frac{1}{25}}$$

$$\epsilon \geq \frac{1}{5}$$

Problem 8

(a) $E(X) = \sum_{i=1}^n (E(X_i)P[X = i])$

$$E(X) = \sum_{i=1}^n (E(X_i)(1-p)^{i-1}p)$$

$$E\left(\frac{1}{X}\right) = \sum_{i=1}^n \left(E\left(\frac{1}{X_i}\right)(1-p)^{i-1}p\right)$$

$$E\left(\frac{1}{X}\right) = p + \sum_{i=2}^n \left(E\left(\frac{1}{X_i}\right)(1-p)^{i-1}p\right)$$

Since the first term itself can be p the rest of the series would over estimate the value p therefore it is a bad estimator.

- (b) Lets say we have n i.i.d. samples. These values can hold a value of 1 or 0. Y is the fraction of samples that equal 1. Using the expected value for each X_i we can see it's p . so each X_i has a value expected value of p and we have n terms so the total sum is np . The fraction is then

$$\frac{np}{n} = p$$

Problem 9

$$Pr[D] = Pr[D|A]P[A] + Pr[D|B]Pr[B] + Pr[D|C]Pr[C]$$

$$Pr[W|D] = \frac{Pr[W \wedge D]}{P[D]} = \frac{Pr[D|W]P[W]}{P[D]}$$

$$Pr[D] = (1/3)(0.4 \times 0.6^3) + (1/3)(.5^4) + (1/3)(0.6 \times 0.4^3)$$

$$Pr[W|D] = \frac{0.4 \times 0.6^3 \times (1/3)}{(1/3)(0.4 \times 0.6^3) + (1/3)(.5^4) + (1/3)(0.6 \times 0.4^3)} = .46$$