.

CSM 70 Discrete Mathematics and Probability Theory

Spring 2016 Random Vars, Expectation, Variance, Distributions, Markov and Chebyshev

Solutions

I. Expectation and Random Variables

Random variable: a function X : Ω → R that assigns a real number to every outcome ω in the probability space.

Expectation: The expectation of a random variable X is defined as

$$E(X) = \sum_{a \in \mathscr{A}} a \times \Pr[X = a]$$

Where the sum is over all possible values taken by the random variable.

Does the random variable always take on the value of its expectation?

No

Make a random variable from the probability space: {2, 3, 6, 7} that half the time is 1 and the other half the time is 0. What function can represent this random variable?

Several answers: x < 5, x%2 = 0, (x = 2 or x = 7), etc.

Given the random variable X defined as taking on the value 1 with probability .25, 2 with probability .5, and 20 with probability .25, what is the expectation of X?

E(X) = .25*1 + .5*2 + .25*20 = 6.25

Linearity of Expectation: For any two random variables X and Y on the same probability space, we have:

$$E(X+Y) = E(X) + E(Y)$$

Also for any constant c:

$$E(cX) = cE(X)$$

Assume we have a biased coin that comes up heads 65% of the time. We also have 4 20-sided dice, each with numbers 1-20. What is the expectation of the sum of the value of the coin (where heads is 1 and tails is 0) and the four dice.

E(coin) = 0.65

E(die) = 10.5

E(4 dice) = 42
E(all) = E(coin) + E(4 dice) = 0.65 + 42 = 42.65

Given the random variable $X = (\text{Summation}(X_i)) + 30$ where X_i is a roll of a die (with i from 1 to 100), what is the expectation of X_1? What is the expectation of $100*X_1 + 30$? What is the expectation of X?

$E(X_1) = ⅙ * 1 + ⅙ * 2 + ⅙ * 3 + ⅙ * 4 + ⅙ * 5 + ⅙ * 6 = 3.5$

$E(100*X_1 + 30) = 100*E(X_1) + 30 = 350 + 30 = 380$

$E(\text{Sum}_i(X_i) + 30) = E(\text{Sum}_i(X_i) + 30) = E(\text{Summ}(X_i)) + 30 = \text{Sum}_i(E(X_i)) + 30 = 100*E(X_1) + 30 = 380$

True or False? $E[X]^4 \leq E[X^4]$

TRUE. First, for an arbitrary random variable Y, we have: $0 \leq E[(Y - E[Y])^2] = E[Y^2] - E[Y]^2$

So $E[Y]^2 \leq E[Y^2]$. Now applying this twice, once for $Y = X$ and once for $Y = X^2$:

$E[X]^4 = (E[X]^2)^2 \leq (E[X^2])^2 \leq E[(X^2)^2] = E[X^4]$

II. Variance

For a random variable X with expectation $E(X) = \mu$, the variance of X is:

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

The square root of Var(X) is called the standard deviation of X

Theorem: For a random variable X with expectation $E(X) = \mu$ and a constant c,

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

$$\text{Var}(cX) = c^2\,\text{Var}(X)$$

Let’s consider the classic problems of flipping coins and rolling dice. Let X be a random variable for the number of coins that land on heads and Y be the value of the die roll.

What is the expected value of X after flipping 3 coins? What is the variance of X?

$E(X) = 0*⅛ + 1*⅜ + 2*⅜ + 3*⅛ = 3/2$

$E(X^2) = 0^2*⅛ + 1^2*⅜ + 2^2*⅜ + 3^2*⅛ = 24/8 = 3$

$E(X)^2 = 9/4 \qquad \text{Var}(X) = 3 - 9/4 = 3/4$

Let Y be the sum of rolling a dice 1 times. What is the expected value of Y?

$E(Y) = (⅙)*(1+2+3+4+5+6) = 7/2$

What is the variance of Y?

$E(Y^2) = [(⅙)(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2)] = 91/6$

$\text{Var}(Y) = E(Y^2) - (E(Y))^2 = 91/6 - (7/2)^2 = 35/12$

Say you’re playing a game with a coin and die, where you flip the coin 3 times and roll the die once. In this game, your score is given by the number of heads that show multiplied with the die result. What is the expected value of your score? What's the variance?

$E(XY) = E(X)E(Y) = 21/4$, since X and Y are independent.
$Var(XY) = E(X^2Y^2) - E(XY)^2 = E(X^2)E(Y^2) - E(X)^2E(Y)^2$
$= 3*(91/6) - (3/2)^2(7/2)^2 = 91/2 - (9/4)(49/4) = 17.9375 = 287/16$

You are at a party with n people where you have prepared a red solo cup labeled with their name. Before handing red cups to your friends, you pick up each cup and put a sticker on it with probability 1/2 (independently of the other cups). Then you hand back the cups according to a uniformly random permutation. Let X be the number of people who get their own cup back AND it has a sticker on it.

a) Compute the expectation E(X).
Define X_i = 1 if the i-th person gets their own cup back and it has a sticker on it
0 otherwise
Hence $E(X) = E(\sum\{i=1^n\}(Xi) = \sum\{i=1^n\} E(X_i)$
$E(X_i) = Pr[X_i = 1] = 1 / 2n$ since the i-th student will get his/her cup with probability 1/n and has a sticker on it with probability 1/2 and stickers are put independently.
Hence $E(X) = n \cdot 1/2n = 1/2$.

b) Compute the variance Var(X)
To calculate Var(X), we need to know $E(X^2)$
$E(X^2) = E(X_1 + X_2 + ... + X_n)^2 = E(\sum\{i, j\}(XiXj)) = \sum\{i,j\} (E(X_iX_j)$ (by linearity of expectation)

Then we consider two cases, either $i = j$ or $i \neq j$.
Hence $\sum \{i,j\} E(X_iX_j) = \sum \{i\} E(X_i^2) + \sum\_ \{i \neq j\} E(X_iX_j)$
$E(X_i^2) = 1/2n$ for all i.
To find $E(X_iX_j)$, we need to calculate $Pr[X_iX_j = 1]$. $Pr[X_iX_j = 1] = Pr[Xi = 1] Pr[Xj = 1 | Xi = 1] = 1/2n \cdot 1/2(n-1)$ since if student i has received his/her own homework, student j has $n - 1$ choices left.
Hence $E(X^2) = n \cdot 1/2n + n \cdot (n - 1) \cdot 1/2n \cdot 1/2(n-1) = 3/4$.
$Var(X) = E(X^2) - E(X)^2 = 3/4 - 1/4 = 1/2$.

True or False? Assume that X is a discrete random variable. If Var(X) = 0, then X is a constant.
TRUE. Let $\mu = E[X]$. By definition, $0 = Var(X) = E[(X - \mu)^2] = \sum \omega \in \Omega\ Pr[\omega](X(\omega) - \mu)^2$ The RHS is the sum of non-negative numbers, so if the sum is 0, each term must be 0. So $Pr[\omega] > 0 \implies (X(\omega) - \mu) 2 = 0 \implies X(\omega) = \mu$. Therefore X is constant (equal to $\mu = E[X]$).

III. Distribution (Geometric, Poisson, Binomial)

Geometric Distribution: Geom(p) Number of trials required to obtain the first success. Each trial has probability of success equal

to p. The probability of the first success happening at trial k is:

$$\Pr(X = k) = (1-p)^{k-1}p, \quad k > 0$$

The expectation of a geometric distribution is:

$$E(X) = \frac{1}{p}$$

Can walk through the derivation of E(X): (Sinho)
The clever way to find the expectation of the geometric distribution uses a method known as the renewal method. E(X) is the expected number of trials until the first success. Suppose we carry out the first trial, and one of two outcomes occurs. With probability p, we obtain a success and we are done (it only took 1 trial until success). With probability 1 − p, we obtain a failure, and we are right back where we started. In the latter case, how many trials do we expect until our first success? The answer is 1 + E(X): we have already used one trial, and we expect E(X) more since nothing has changed from our original situation (the geometric distribution is memoryless).
Hence E(X) = p · 1 + (1 − p) · (1 + E(X))

Binomial Distribution: Bin(n, p)
Number of successes when we do n independent trials. Each trial has a probability p of success. The probability of having k successes:

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

The expectation of a binomial distribution is:

$$E(X) = np$$

Can walk through the derivation of E(X): (Sinho)
We would have to compute this sum:

$$E(X) = \sum_k k \Pr(X = k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

Instead of doing that just use Bernoulli variables:

$$X = X_1 + \cdots + X_n$$

And now use linearity of expectation:

$$\begin{aligned} E(X) &= E(X_1 + \cdots + X_n) \\ &= E(X_1) + \cdots + E(X_n) \end{aligned}$$

Since the probability of a success happening at each step is p, and there are n steps, we are just summing p n times.

Poisson Distribution: Pois(λ)
This is an approximation to the binomial distribution. Let the number of trials approach

infinity, let the probability of success approach 0, such that E(X) = $np = \lambda$. This is an accepted model for "rare events". The probability of having k successes:

$$\Pr(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

The expectation of a poisson distribution is:

$$E(X) = \lambda$$

Can walk through the derivation of P(X): (Sinho)

$$\begin{aligned}\Pr(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \\ &\approx \frac{n^k p^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \\ &\approx \frac{\lambda^k e^{-\lambda}}{k!}\end{aligned}$$

Can walk through derivation of E(X): (Sinho)

$$\begin{aligned}E(X) &= \sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} \\ &= \sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} \\ &= e^{-\lambda}\lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= e^{-\lambda}\lambda \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda}\lambda e^{\lambda} \\ &= \lambda\end{aligned}$$

You are Eve, and as usual, you are trying to break RSA. You are trying to guess the factorization of N, from Bob's public key. You know that N is approximately 1,000,000,000,000. To find the primes p and q, you decide to try random numbers from 2 to 1,000,000 ≈ sqrt(N), and see if they divide N.

To do this, you roll a 999,999-sided die to choose the number, and see if it divides N using your calculator, which takes five seconds. Of course, there will be one number in this range that does divide N—namely, the smaller of p and q.

1. What kind of distribution would you use to model this?
   Geometric — probability of success each time is p = 1 / 999,999

2. What is the expected *amount of time* until you guess the correct answer, if it takes five seconds per guess (you only have a calculator)? Answer in days.
   E[x] = 1/p = 999,999 tries
   (999,999 * 5 sec) / (60 sec/min) / (60 min/hr) / (24 hr/day) ≈ 57.9 days

3. What is the variance in the amount of time? (Answer in seconds, approximately.)
   Var(x) = (1-p)/p^2 = (1-1/999,999) / (999,999)^2 = 999,999 * 999,998 ≈ 1 trillion tries
   When we scale a variable, we scale the variance by the square of that factor, so Var(x) = 25 trillion sec^2

Now you are trying to guess the 6-digit factorization digit by digit. Let's assume that when you finish putting these digits together, you can figure out how many digits you got right. Use zeros for blank spaces. For example, to guess 25, you would put 000025

1. What kind of distribution would you use to model this?
   Binomial, since this is multiple independent trials that can either succeed or fail.

2. What is the probability that you get exactly 4 digits right?
   (6 choose 4) * $(1/10)^4$ * $(9/10)^2$

3. What is the probability that you get less than 3 correct?
   (6 choose 2) * $(1/10)^2$ * $(9/10)^4$ + (6 choose 1) * (1/10) * $(9/10)^5$

You are Alice, and you have a high-quality RSA-based security system. However, Eve is often successful at hacking your system. You know that the number of security breaches averages 3 a day, but varies greatly.

1. What kind of distribution would you use to model this?
   Poisson! That's what we use to model the probably frequencies of rare events.

2. What is the probability you experience exactly seven attacks tomorrow? At least seven (no need to simplify your answer)?

$$Pr[X=7] = \frac{\lambda^7}{7!}e^{-\lambda} = \frac{3^7}{7!}e^{-3} \approx 0.0216$$

$$Pr[X \geq 7] = \sum_{i=7}^{\infty} Pr[X=i] = \sum_{i=7}^{\infty} \frac{3^i}{i!}e^{-3}$$

3. What is the probability that, on some day in April, you experience exactly six attacks?

$$Pr[X=6] = \frac{3^6}{6!}e^{-3} \approx 0.0504$$

$$1-(1-0.0504)^{30} \approx 0.788 = 78.8\%$$

IV. Markov and Chebyshev

Markov's Inequality
For a non-negative random variable X with expectation E(X) = μ, and any α > 0:

$$\Pr[X \geq \alpha] \leq \frac{\mathrm{E}(X)}{\alpha}$$

Proof as part of lesson plan:

$$\begin{aligned}\mathrm{E}(X) &= \textstyle\sum_a a \times \Pr[X = a] \\ &\geq \textstyle\sum_{a \geq \alpha} a \times \Pr[X = a] \\ &\geq \alpha \textstyle\sum_{a \geq \alpha} \Pr[X = a] \\ &= \alpha \Pr[X \geq \alpha].\end{aligned}$$

Chebyshev's Inequality
For a random variable X with expectation E(X) = μ, and any α > 0:

$$\Pr[|X - \mu| \geq \alpha] \leq \frac{\mathrm{Var}(X)}{\alpha^2}$$

Use Markov's to prove Chebyshev's Inequality:
Define the random variable $Y = (X - \mu)^2$. Note that $E(Y) = E((X - \mu)^2) = \mathrm{Var}(X)$
Also, notice that the event that we are interested in, $|X - \mu| \geq \alpha$ is exactly the same as the event $Y = (X - \mu)^2 \geq \alpha^2$.
Therefore, $\Pr[|X - \mu| \geq \alpha] = \Pr[Y \geq \alpha^2]$. Moreover, Y is obviously non-negative, so we can apply Markov's inequality to it to get

$$\Pr[Y \geq \alpha^2] \leq \frac{\mathrm{E}(Y)}{\alpha^2} = \frac{\mathrm{Var}(X)}{\alpha^2}.$$

A random variable X always takes on values greater than -60. Find the best bound possible for Pr[X>= -10] when E[X] = -35.

Solutions: Take the Markov for Y which is X + 60. Notice that now, Y is positive always, so we can apply Markov to it, while before, X had the possibility of being negative. Taking the Markov Bound of this, we get that Pr[Y>=50] <= 25/50 = ½. The 50 came from adding 60 to -10, while the E[Y] is E[X] + 60.

Consider a coin that comes up with head with probability 0.2 . Let us toss it n times. Use Markov's to bound the probability of getting 80 percent heads.

Let X be the amount of heads in n flips.

E[X] = .2n, by the binomial distribution.

So, P(X >= .8n) <= .2n / .8n, which is just .25.

As we all know, Berkeley squirrels are extremely fat and cute. The average squirrel is 40% body fat. The standard deviation of body fat is 5%. Provide an upper bound on the probability that a randomly trapped squirrel is either too skinny or too fat? A skinny squirrel has less than 27.5% body fat, and an obese squirrel has more than 52.5% body fat?

We use Chebyshev's inequality. We are looking for the probability we fall within 2.5 standard deviations of the mean. By Chebyshev's inequality, the probability we are within this range is 1/(2.5)^2, or 4/25 = 0.16.
If we were to use Markov's inequality, we would probabilities over 1, which yields a non-helpful value.