INEQUALITIES, DISTRIBUTIONS, CONTINUOUS PROBABILITY, CONDITIONAL EXPECTATION 9

COMPUTER SCIENCE MENTORS 70

November 7 to November 11, 2016

1 Distributions

1.1 Introduction

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:

$$\mathrm{P}[X = k] = (1-p)^{k-1} * p, k > 0$$

The expectation of a geometric distribution is:

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

The variance of a geometric distribution is:

$$Var(X) = \frac{1-p}{p^2}$$

Solution: Derivation of $\mathrm{E}(X)$: The clever way to find the expectation of the geometric distribution uses a method known as the renewal method. $\mathrm{E}(X)$ 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:

$$\text{P}[X = k] = \binom{n}{k} * p^k * (1-p)^{n-k}$$

The expectation of a binomial distribution is:

$$\text{E}(X) = np$$

The variance of a binomial distribution is:

$$Var(X) = np(1-p)$$

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

$$\text{E}(X) = \sum_k k * \text{P}[X = k] = \sum_{k=0}^{n} k * \binom{n}{k} * p^k * (1-p)^{n-k}$$

Instead of doing that just use Bernoulli variables:

$$X = X_1 + \ldots + X_n$$

And now use linearity of expectation:

$$\text{E}(X) = \text{E}(X_1 + \ldots + X_n) = \text{E}(X_1) + \ldots + \text{E}(X_n)$$

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 $\mathrm{E}(X) = np = \lambda$. This is an accepted model for rare events. The probability of having k successes:

$$\mathrm{P}[X = k] = \frac{e^{-\lambda} * \lambda^k}{k!}$$

The expectation of a poisson distribution is:

$$\mathrm{E}(X) = \lambda$$

The variance of a poisson distribution is:

$$Var(X) = \lambda$$

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

$$\begin{aligned}
\mathrm{P}[X = k] &= \binom{n}{k} * p^k * (1-p)^{n-k} && (1)\\
&= \frac{n!}{k! * (n-k)!} * p^k * (1-p)^{n-k} && (2)\\
&\approx \frac{n^k * p^k}{k!} * (1 - \frac{\lambda}{n})^{n-k} && (3)\\
&\approx \frac{\lambda^k * e^{-\lambda}}{k!} && (4)
\end{aligned}$$

To get from equation (2) to (3) we make use of the assumption that

$$n$$

is large for the approximation

$$\frac{n!}{(n-k)!} \approx n^k$$

and to get from (3) to (4) we calculate the limit of $(1 - \frac{\lambda}{n})^{n-k}$ as n approahces ∞.

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

$$= e^{-\lambda} * \lambda * \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$$
$$= e^{-\lambda} * \lambda * \sum_{k=1}^{\infty} \frac{\lambda^{k}}{k!}$$
$$= e^{-\lambda} * \lambda * e^{\lambda}$$
$$= \lambda$$

We note that this expectation is exactly what we want it to be (it matches the expectation of the binomial).

1.2 Questions

1. You are Eve, and as usual, you are trying to break RSA. You are trying to guess the factorization of N, from Bobs 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 $1,000,000 \approx \sqrt{N}$, 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 Nnamely, the smaller of p and q.

   (a) What kind of distribution would you use to model this?

   Solution: Geometric probability of success each time is $p = \frac{1}{999{,}999}$

   (b) 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.

   Solution:
   $$\mathrm{E}(x) = \frac{1}{p} = 999,999 \text{ tries}$$
   $$(999,999 * 5 \text{ sec}) * \frac{1 \text{ min}}{60 \text{ sec}} * \frac{1 \text{ hr}}{60 \text{ min}} * \frac{1 \text{ day}}{24 \text{ hr}} \approx 57.9 \text{ days}$$

2. Now you are trying to guess the 6-digit factorization digit by digit. Lets 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

   (a) What kind of distribution would you use to model this?

   Solution: Binomial, since this is multiple independent trials that can either succeed or fail.

   (b) What is the probability that you get exactly 4 digits right?

   Solution: $\binom{6}{4} * \frac{1}{10}^4 * \frac{9}{10}^2$

   (c) What is the probability that you get less than 3 correct?

   Solution: $\binom{6}{2} * \frac{1}{10}^2 * \frac{9}{10}^4 + \binom{6}{1} * \frac{1}{10} * \frac{9}{10}^5 + \binom{6}{0} * \frac{1}{10}^0 * \frac{9}{10}^6$

3. 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.

   (a) What kind of distribution would you use to model this?

   Solution: Poisson! That's what we use to model the probably frequencies of rare events.

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

   Solution:

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

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

   (c) What is the probability that, on some day in April, you experience exactly six attacks?

Solution:

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

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

2 Variance

2.1 Introduction

Definition: For a random variable X, expectation $\mathrm{E}(X) =$

$$\sum_a a * Pr[X = A]$$

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

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

The square root of Var(X) is called the standard deviation of X and is often denoted σ.

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

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

$$\mathrm{Var}(cX) = c^2 * \mathrm{Var}(X)$$

2.2 Questions

1. 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.

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

Solution:

$$\mathrm{E}(X) = 0 * \frac{1}{8} + 1 * \frac{3}{8} + 2 * \frac{3}{8} + 3 * \frac{1}{8} = \frac{3}{2}$$

$$\mathrm{E}(X^2) = 0^2 * \frac{1}{8} + 1^2 * \frac{3}{8} + 2^2 * \frac{3}{8} + 3^2 * \frac{1}{8} = \frac{24}{8} = 3$$

$$\mathrm{E}(X)^2 = \frac{9}{4}$$
$$\mathrm{Var}(X) = 3 - \frac{9}{4} = \frac{3}{4}$$

(b) Let Y be the sum of rolling a dice 1 time. What is the expected value of Y?

Solution: $\mathrm{E}(Y) = \frac{1}{6} * (1 + 2 + 3 + 4 + 5 + 6) = \frac{7}{2}$

(c) What is the variance of Y?

Solution: $\mathrm{E}(Y^2) = [\frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2)] = \frac{91}{6}$ $\mathrm{Var}(Y) = \mathrm{E}(Y^2) - (\mathrm{E}(Y))^2 = \frac{91}{6}\frac{7}{2}^2 = \frac{35}{12}$

2. 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? Whats the variance?

Solution: $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y) = \frac{21}{4}$ since X and Y are independent. $\mathrm{Var}(XY) = \mathrm{E}(X^2Y^2) - \mathrm{E}(XY)^2 = \mathrm{E}(X^2)\mathrm{E}(Y^2) - \mathrm{E}(X)^2\mathrm{E}(Y)^2 = 3 * \frac{91}{6} - \frac{3}{2}^2 * \frac{7}{2}^2 = \frac{91}{2} - \frac{9}{4} * \frac{49}{4} = 17.9375 = \frac{287}{16}$

3. 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 $\frac{1}{2}$ (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 $\mathrm{E}(X)$.

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

(b) Compute the variance $\mathrm{Var}(X)$

Solution: To calculate $\mathrm{Var}(X)$, we need to know $\mathrm{E}(X^2)$

$$\mathrm{E}(X^2) = \mathrm{E}(X_1 + X_2 + \ldots + X_n)^2 = \mathrm{E}(\sum_{i,j}(X_i * X_j)) = \sum_{i,j}(\mathrm{E}(X_i * X_j)$$

(by linearity of expectation)

Then we consider two cases, either $i = j$ or $i \neq j$. Hence

$$\sum_{i,j} \mathrm{E}(X_i * X_j) = \sum_{i} \mathrm{E}(X_i^2) + \sum_{i \neq j} \mathrm{E}(X_i * X_j)$$

$E(X_i^2) = \frac{1}{2n}$ for all i. To find $\mathrm{E}(X_i * X_j)$, we need to calculate $\mathrm{P}[X_i X_j = 1]$. $\mathrm{P}[X_i * X_j = 1] = \mathrm{P}[X_i = 1]\mathrm{P}[X_j = 1 | X_i = 1] = \frac{1}{2n} * \frac{1}{2*(n-1)}$ since if student i has received his/her own cup, student j has $n - 1$ choices left. Hence

$$\mathrm{E}(X^2) = n * \frac{1}{2n} + n * (n-1) * \frac{1}{2n} * \frac{1}{2 * (n-1)} = \frac{3}{4}$$

$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$.

4. a. Prove that for independent random variables X and Y, $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$.

Solution:

$$\begin{aligned}\text{Var}(X+Y) &= \text{E}((X+Y)^2) - \text{E}(X+Y)^2 \\ &= \text{E}(X^2) + \text{E}(Y^2) + 2 * \text{E}(XY) - (\text{E}(X) + \text{E}(Y))^2 \\ &= (\text{E}(X^2) - \text{E}(X)^2) + (\text{E}(Y^2) - \text{E}(Y)^2) + 2 * \text{E}(XY) - \text{E}(X) * \text{E}(Y) \\ &= \text{Var}(X) + \text{Var}(Y) + 2 * (\text{E}(XY) - \text{E}(X)\text{E}(Y))\end{aligned}$$

b. Is the above result true for non-independent random variables? Prove or give a counterexample.

Solution: No! One simple counterexample is $X = Y$. Then

$$\begin{aligned}\text{Var}(X+Y) &= \text{Var}(2x) \\ &= \text{E}((2X)^2) - \text{E}(2X)^2 \\ &= \text{E}(4X^2) - (2 * \text{E}(X))^2 \\ &= 4\text{E}(X^2) - 4(\text{E}(X)^2) \\ &= 4(\text{E}(X^2) - \text{E}(X)^2) \\ &= 4 * \text{Var}(X)\end{aligned}$$

5. Consider the random variable $X = X_1 + \ldots + X_n$, where X_i equals i with probability $\frac{1}{i}$ and 0 otherwise.

(a) What is the variance of X? (Assume that X_i and X_j are independent for $i \neq j$)

Solution:

$$\begin{aligned}\text{Var}(X) &= \text{Var}(X_1) + \ldots + \text{Var}(X_n) \\ \text{E}(X_i^2) &= \text{P}[X_i = i] * i^2 + \text{P}[X_i = 0] * 0^2 = \frac{1}{i} * i^2 + 0 = i \\ (\text{E}(X_i))^2 &= (\text{P}[X_i = i] * i + \text{P}[X_i = 0] * 0 = \frac{1}{i} * i + 0 = 1 \\ \text{Var}(X_i) &= \text{E}(X_i^2) - (\text{E}(X_i))^2 = i - 1\end{aligned}$$

Recall,

$$\sum_{k=1}^{n} = \frac{n * (n+1)}{2}$$

$$
\begin{aligned}
\mathrm{Var}(X) &= \sum_i \mathrm{Var}(X_i) \\
&= \sum_i i - 1 \\
&= -n + \sum_i i \\
&= -n + \frac{n*(n+1)}{2} \\
&= \frac{n*(n+1) - 2n}{2} \\
&= \frac{n^2 + n - 2n}{2} \\
&= \frac{n^2 - n}{2} \\
&= \frac{n*(n-1)}{2}
\end{aligned}
$$

(b) For what value of n does $\mathrm{E}(X) = \mathrm{Var}(X)$?

Solution:

$$
\begin{aligned}
\mathrm{E}(X_i) &= \mathrm{P}[X_i = i] * i + 0 \\
&= \frac{1}{i} * i = 1
\end{aligned}
$$

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

$$
\begin{aligned}
n = \frac{n*(n-1)}{2} &\to 1 = \frac{n-1}{2} \\
&\to 2 = n - 1 \\
&\to 3 = n
\end{aligned}
$$

(c) For what value of n does $E(X) = SD(X) * \sqrt{2} + 100$?

Solution: $\mathrm{E}(X) = n, SD(X) = \sqrt{\mathrm{Var}(X)}$

$$\begin{aligned} n &= \sqrt{\frac{n*(n-1)}{2}} * \sqrt{2} + 100 \\ n &= \sqrt{n*(n-1)} + 100 \\ (n-100)^2 &= n*(n-1) \\ n^2 - 200n + 10000 &= n^2 - n \\ 10000 &= 199n \\ n &= \frac{10000}{199} \end{aligned}$$

6. An urn contains n balls numbered $1, 2, \ldots, n$. We remove k balls at random (without replacement) and add up their numbers. Find the mean and variance of the total.

Solution: The required total is $T = \sum_{i=1}^{k} X_i$, where X_i is the number shown on the ith ball. Hence $\mathrm{E}(T) = k * \mathrm{E}(X_1) = \frac{1}{2} * k * (n+1)$. Now calculate:

$$\begin{aligned} \mathrm{E}((\sum_{i=1}^{k} X_i)^2) &= k\mathrm{E}(X_1^2) + k*(k-1)*\mathrm{E}(X_1 * X_2) \\ &= \frac{k}{n}\sum_{1}^{n} j^2 + \frac{k*(k-1)}{n*(n-1)} * 2 * \sum_{i>j} i*j \\ &= \frac{k}{n}(\frac{1}{3}*n*(n+1)*(n+2) - \frac{1}{2}*n*(n+1)) \\ &+ \frac{k*(k-1)}{n*(n-1)} * \sum_{j=1}^{n} j*(n*(n+1) - j*(j+1)) \\ &= \frac{1}{6}*k*(n+1)*(2n+1) + \frac{1}{12}*k*(k-1)*(3n+2)*(n+1) \end{aligned}$$

Hence,

$$\begin{aligned} \mathrm{Var}(T) &= k(n+1)(\frac{1}{6}k(n+1)(2n+1) + \frac{1}{12}k(k-1)(3n+2)(n+1) - \frac{1}{4}k(n+1)) \\ &= \frac{1}{12}(n+1)k(n-k) \end{aligned}$$

3 Conditional Expectation

3.1 Introduction

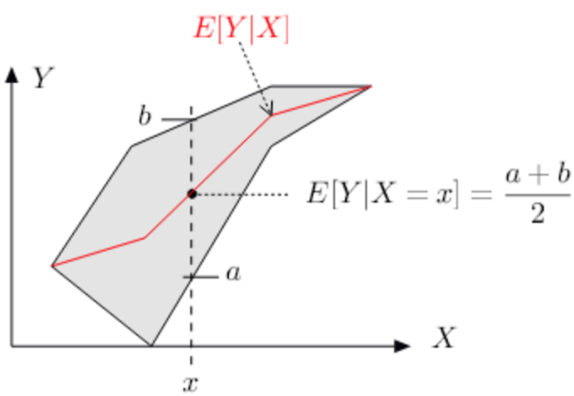
The **conditional expectation** of Y given X is defined by

$$E[Y|X = x] = \sum_y y \cdot P[Y = y|X = x] = \sum_y y \cdot \frac{P[X = x, Y = y]}{P[X = x]}$$

Properties of Conditional Expectation

$$\begin{aligned} E(a|Y) &= a \\ E(aX + bZ|Y) &= a \cdot E(X|Y) + b \cdot E(Z|Y) \\ E(X|Y) &\geq 0 \text{ if } X \geq 0 \\ E(X|Y) &= E(X) \text{ if } X, Y \text{ independent} \\ E(E(X|Y)) &= E(X) \end{aligned}$$

Solution: Here is a picture that shows that conditioning creates a new random variable with a new distribution. Figure 9 of note 26 does so.



3.2 Questions

1. Prove $E(E(Y|X)) = E(Y)$

Solution:

$$E(E(Y|X)) = \sum_x E(Y|X = x) \cdot P[X = x]$$

$$
\begin{aligned}
&= \sum_x \left(\sum_y y \cdot \mathrm{P}[Y = y|X = x]\right) \cdot \mathrm{P}[X = x] \\
&= \sum_y y \cdot \sum_x y \cdot \mathrm{P}[X = x|Y = y]) \cdot \mathrm{P}[Y = y] \\
&= \sum_y y \cdot \mathrm{P}[Y = y] \cdot \sum_x \mathrm{P}[X = x|Y = y]) \\
&= \sum_y y \cdot \mathrm{P}[Y = y] = E[Y]
\end{aligned}
$$

2. Prove $\mathrm{E}(h(X) \cdot Y|X) = h(X) \cdot \mathrm{E}(Y|X)$

Solution:

$$
\begin{aligned}
\mathrm{E}(h(X) \cdot Y|X) &= \sum_y h(X) \cdot y \cdot \mathrm{P}[Y = y|X] \\
&= h(X) \sum_y y \cdot \mathrm{P}[Y = y|X] \\
&= h(X) \cdot \mathrm{E}[Y|X]
\end{aligned}
$$