CSE 312: Foundations of Computing II
Quiz Section #6: Variance, Independence, Zoo of Discrete Random Variables

Review/Mini-Lecture/Main Theorems and Concepts From Lecture

Variance: Let X be a random variable and $\mu = E[X]$. The variance of X is defined to be

$$Var(X) = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

Notice that since this is an expectation of a _____________ random variable ($(X-\mu)^2$), variance is always _____________. With some algebra, we can simplify this to $Var(X) = E[X^2] - E^2[X]$.

Independence: Random variables X and Y are independent, written $X \perp Y$, iff

____

In this case, we have $E[XY] = E[X]E[Y]$ (the converse is not necessarily true).

i.i.d. (independent and identically distributed): Random variables $X_1, \dots, X_n$ are i.i.d. (or iid) if they are ___________ and have the same ___________________________________.

Property of Variance: Let $a, b \in \mathbb{R}$ and X a random variable. Then,

$$Var(aX + b) = \_\_\_\_\_\_\_\_\_$$

Linearity of Variance: If $X \perp Y$, $Var(X+Y) = Var(X) + Var(Y)$. Linearity of variance depends on independence, whereas linearity of expectation always holds. Note that this combined with the above show that $\forall a, b, c \in \mathbb{R}$ and if $X \perp Y$,

$$Var(aX + bY + c) = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

Zoo of Discrete Random Variables

Uniform: $X \sim Unif(a, b)$ if X has the following probability mass function:

$$p_X(k) = \frac{1}{b-a+1}, \qquad k = a, \dots, b$$

$E[X] = \frac{a+b}{2}$ and $Var(X) = \frac{(b-a)(b-a+2)}{12}$. This represents each integer from $[a, b]$ to be equally likely. For example, a single roll of a fair die is $Unif(1,6)$.

Bernoulli (or indicator): $X \sim Ber(p)$ if X has the following probability mass function:

$$p_X(k) = \begin{cases} p, & k = 1 \\ 1-p, & k = 0 \end{cases}$$

$E[X] = p$ and $Var(X) = p(1-p)$. An example of a Bernoulli r.v. is one flip of a coin with $P(head) = p$. By a clever trick, we can write

$$p_X(k) = p^k(1-p)^{1-k}, \qquad k = 0,1$$

Binomial: $X \sim Bin(n, p)$ if X is the sum of iid $Ber(p)$ random variables, and has pmf

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0,1,\dots,n$$

$E[X] = np$ and $Var(X) = np(1-p)$. An example of a Binomial r.v. is the number of heads in n independent flips of a coin with $P(head) = p$. Note that $Bin(1,p) \equiv Ber(p)$. As $n \to \infty$ and $p \to 0$, with $np = \lambda$, then $Bin(n,p) \to Poi(\lambda)$. If $X_1, \dots, X_n$ are independent Binomial r.v.'s, where $X_i \sim Bin(N_i, p)$, then $X = X_1 + \cdots + X_n \sim Bin(N_1 + \cdots + N_n, p)$.

Geometric: $X \sim Geo(p)$ if X has the following probability mass function:

$$p_X(k) = (1-p)^{k-1}p, \qquad k = 1,2,\dots$$

$E[X] = \frac{1}{p}$ and $Var(X) = \frac{1-p}{p^2}$. An example of a Geometric r.v. is the number of independent coin flips up to and including the first head, where $P(head) = p$.

Negative Binomial: $X \sim NegBin(r, p)$ if X is the sum of iid Geometric random variables, and has pmf

$$p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \qquad k = r, r+1, \dots$$

$E[X] = \frac{r}{p}$ and $Var(X) = \frac{r(1-p)}{p^2}$. An example of a Negative Binomial r.v. is the number of independent coin flips up to an including the r^{th} head, where $P(head) = p$. If $X_1, \dots, X_n$ are independent Negative Binomial r.v.'s, where $X_i \sim NegBin(r_i, p)$, then $X = X_1 + \cdots + X_n \sim NegBin(r_1 + \cdots + r_n, p)$.

Poisson: $X \sim Poi(\lambda)$ if X has the following probability mass function:

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0,1,\dots$$

$E[X] = \lambda$ and $Var(X) = \lambda$. An example of a Poisson r.v. is the number of people being born in a minute, where λ is the average rate per unit time. If $X_1, \dots, X_n$ are independent Poisson r.v.'s, where $X_i \sim Poi(\lambda_i)$, then $X = X_1 + \cdots + X_n \sim Poi(\lambda_1 + \cdots + \lambda_n)$.

Hypergeometric: $X \sim HypGeo(N, K, n)$ if X has the following probability mass function:

$$p_X(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \qquad k = \max\{0, n+K-N\}, \dots, \min\{K, n\}$$

$E[X] = n\frac{K}{N}$. This represents the number of successes drawn, when n items are drawn from a bag with N items (K of which are successes, and $N-K$ failures) without replacement. If we did this with replacement, then this scenario would be represented as $Bin\left(n, \frac{K}{N}\right)$.

Section #6 Review

Exercises

1. Suppose I am fishing in a pond with B blue fish, R red fish, and G green fish, where $B + R + G = N$. For each of the following scenarios: identify the most appropriate distribution (with parameter(s)):

a) how many of the next 10 fish I catch are blue, if I catch and release

b) how many fish I had to catch until my first green fish, if I catch and release

c) how many red fish I catch in the next five minutes, if I catch on average r red fish per minute

d) whether or not my next fish is blue

e) how many of the next 10 fish I catch are blue, if I do not release the fish back to the pond after each catch

f) how many fish I have to catch until I catch three red fish, if I catch and release

2. Suppose I have $Y_1, \dots, Y_n$ iid with $E[Y_i] = \mu$ and $Var(Y_i) = \sigma^2$, and let $Y = \frac{1}{n}\sum_{i=1}^{n} iY_i$. What is $E[Y]$ and $Var(Y)$? Recall that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ and $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.

3. Is the following statement true or false? If $E[XY] = E[X]E[Y]$, then $X \perp Y$. If it is true, prove it. If not, provide a counterexample.

4. Suppose we roll two fair 5-sided dice independently. Let X be the value of the first die, Y be the value of the second die, $Z = X + Y$ be their sum, $U = \min\{X, Y\}$ and $V = \max\{X, Y\}$.

a) Find $p_U(u)$.

b) Find $E[U]$.

c) Find $E[Z]$.

d) Find $E[UV]$.

e) Find $Var(U+V)$.

5. Suppose X has the following probability mass function:

$$p_X(x) = \begin{cases} c, & x = 0 \\ 2c, & x = \frac{\pi}{2} \\ c, & x = \pi \\ 0, & otherwise \end{cases}$$

a) Suppose $Y_1 = \sin(X)$. Find $E[Y_1^2]$.

b) Suppose $Y_2 = \cos(X)$. Find $E[Y_2^2]$.

c) Suppose $Y = Y_1^2 + Y_2^2 = \sin^2(X) + \cos^2(X)$. Before any calculation, what do you think $E[Y]$ should be? Find $E[Y]$, and see if your hypothesis was correct. (Recall for any real number x, $\sin^2(x) + \cos^2(x) = 1$).

d) Let W be any discrete random variable with probability mass function $p_W(w)$. Then, $E[\sin^2(W) + \cos^2(W)] = 1$. Is this statement always true? If so, prove it. If not, give a counterexample by giving a probability mass function for a discrete random variable W for which the statement is false.

6. If electricity power failures occur according to a Poisson distribution with an average of 3 failures every twenty weeks, calculate the probability that there will be more than one failure during a particular week.

7. A company makes electric motors. The probability an electric motor is defective is 0.01, independent of other motors made. What is the probability that a sample of 300 electric motors will contain exactly 5 defective motors? Do it first exactly, then approximate it with the Poisson. How good was the approximation?

8. An average page in a book contains one typo. What is the probability that there are exactly 8 typos in a given 10-page chapter, using the Poisson model?

Cool puzzles from earlier topics

9. A plane has 100 seats and 100 passengers. The first person to get on the plane lost his ticket and doesn't know his assigned seat, so he picks a seat uniformly at random to sit in. Every remaining person knows their seat, so if it is available they sit in it, and if it is unavailable they pick a uniform random remaining seat. What is the probability the last person to get on gets to sit in his own seat?

10. Suppose you're on a game show, and you're given the choice of three doors. Behind one door is a car, behind the others, goats. You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say number 3, which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

11. You flip a fair coin independently and count the number of flips until the first tail, including that tail flip in the count. If the count is n, you receive 2^n dollars. What is the expected amount you will receive? How much would you be willing to pay at the start to play this game?