CENG 551 Probability and Stochastic Processes for Engineers

Week # 5

Random Variables: Geometric RV’s

Dr. Deniz Ozdemir

Deniz.ozdemir@yasar.edu.tr

Web: dozdemir.yasar.edu.tr

Annoncements

- No class on October 30 (next week)
  - Recitation?

Plan for the Session

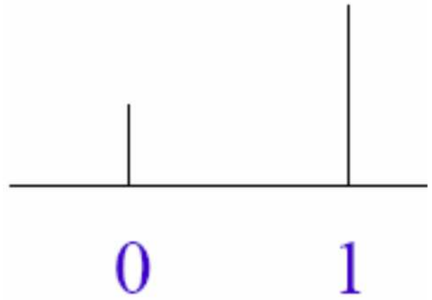
- Geometric Random Variables

- Joint PMF of two random variables
  - Independence revisited

Selected Discrete Distributions:

- Single Coin flip P(heads/success) = p
  - *Probability of success*

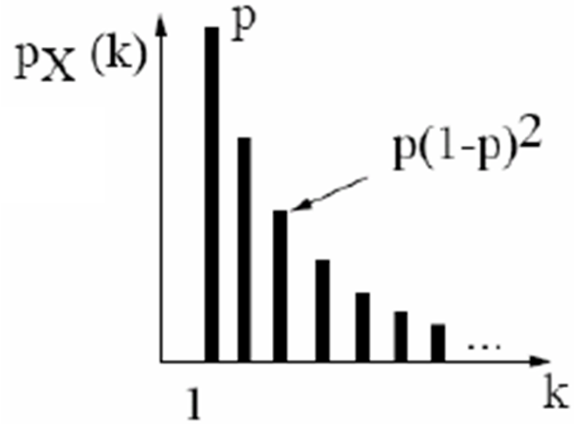
$$f(x) = P(X = x) = \begin{cases} p & \text{if } x = 1 \text{ (success)} \\ 1 - p & \text{if } x = 0 \text{ (failure)} \end{cases}$$



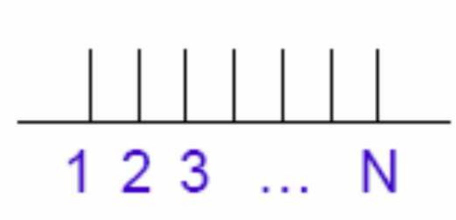
  - *"Bernoulli Distribution"*
- Multiple coin flips
  - x successes out of n trials
  - $f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ for x = 1,…, n
  - *"Binomial distribution"*

Selected Discrete Distributions:

- Number of failures before the first success
  - $f(x)=P(X=x) = (1\text{-}p)^{x\text{-}1}p, \qquad x=1,2,..$
  - *"Geometric Distribution"*



- N equally likely events
  - $f(x)=P(X=x) = 1/N, \qquad x=1,2,..$
  - *"Uniform distribution"*



Properties of Bernoulli Distribution

- PMF: $f(x) = P(X = x) = \begin{cases} p & \text{if x} = 1 \text{ (success)} \\ 1 - p & \text{if x} = 0 \text{ (failure)} \end{cases}$
- Expectation:
  - $E[X] = p$
  - $E[X^k] = p$
- Variance:
  - $E(X^2) - (E[X])^2 = p - p^2 = p(1-p)$

Properties of Binomial Distribution

- PMF: $f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$
- Expectation:
  - $E[X] = np$
  - $E[X^k] = pnE[Y+1]^{k-1}$ where $Y \sim \text{Bin}(n-1, p)$
- Variance:
  - $\text{E}(X^2) - (\text{E}[X])^2 = pn[(n-1)p] - n^2p^2 = np(1-p)$

Useful things to remember:

- Stirling Approximation: $n! \sim n^{n+1/2} e^{-n} \sqrt{2\pi}$

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}$$

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

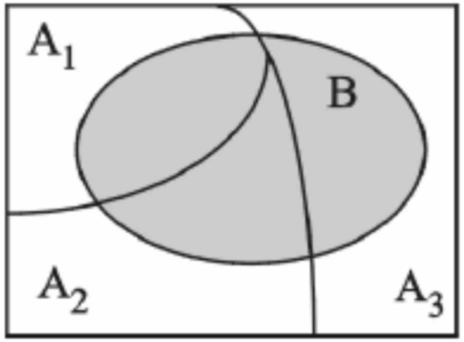
$$\sum_{k=0}^{n} x^k = \frac{1-x^{k+1}}{1-x}$$

$$\sum_{k=0}^{\infty} x^k = \begin{cases} \infty & \text{if } x \geq 1 \\ \dfrac{1}{1-x} & \text{if } x < 1 \end{cases}$$

Properties of Geometric Distribution

- PMF: $f(x)=P(X=x) = (1-p)^{x-1}p, \quad x=1,2,..$
- Expectation:
  - $$E[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \sum_{k=1}^{\infty} kpq^{k-1}$$
    $$= p\sum_{k=1}^{\infty} \underbrace{kq^{k-1}}_{\frac{dq^k}{dq}} = p\frac{d}{dq}\sum_{k=1}^{\infty} q^k$$
    $$= p\frac{d}{dq}\left(\frac{1}{1-q}-1\right) = \frac{p}{(1-q)^2} = \frac{1}{p}$$
  - $E[X^2] =$
    $$\sum_{k=1}^{\infty} k^2 p(1-p)^{k-1} = \sum_{k=1}^{\infty} k^2 pq^{k-1}$$
    $$= p\sum_{k=1}^{\infty} \underbrace{k^2 q^{k-1}}_{\frac{dkq^k}{dq}} = p\frac{d}{dq}\sum_{k=1}^{\infty} kq^k$$
    $$= p\frac{d}{dq}\sum_{k=1}^{\infty} \frac{kq^{k-1}p}{p/q} = p\frac{d}{dq}\left(\frac{q}{1-q}E(X)\right)$$
    $$= p\frac{d}{dq}\left(\frac{q}{(1-q)^2}\right) = p\left(\frac{(1-q)^2+2q(1-q)}{(1-q)^4}\right)$$
    $$= p\left(\frac{1}{p^2}+\frac{2(1-p)}{p^3}\right) = \frac{2}{p^2}-\frac{1}{p}$$
- Variance:
  - $E(X^2) - (E[X])^2 = 1/p^2 - 1/p$

Properties of Geometric Distribution



- Total Expectation Theorem
- Partition of sample space into disjoint events: $A_1, A_2, \ldots, A_n$

$$\mathbf{P}(B) = \mathbf{P}(A_1)\mathbf{P}(B|A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B|A_n)$$

$$\mathbf{E}[X] = \mathbf{P}(A_1)\mathbf{E}[X|A_1] + \cdots + \mathbf{P}(A_n)\mathbf{E}[X|A_n]$$

- Geometric example: $A_1 : \{X = 1\},\ A_2 : \{X > 1\}$
  - $E(X)$

$= E(X|X=1)P(X=1) + E(X|X>1)P(X>1)$

$= p\,[1] + (1-p)\,[E(X)+1] = E(X) + 1 - pE(X)$

- Solve to get: $E(X) = 1/p$

Exercise:

- Two coins are simultaneously tossed until one of them comes up a head and the other a tail. The first coin comes up a head with probability p and the second with probability q. All tosses are assumed independent.
  - Find the PMF, the expected value, and the variance of the number of tosses.

Let X be the # of tosses until the game is over. Note that X is geometric with probability of success

$$\mathbf{P}\big(\{HT,TH\}\big) = p(1-q) + q(1-p),$$

we obtain

$$p_X(k) = \big(1 - p(1-q) - q(1-p)\big)^{k-1}\big(p(1-q) + q(1-p)\big), \qquad k = 1, 2, \ldots$$

Therefore

$$\mathbf{E}[X] = \frac{1}{p(1-q) + q(1-p)}$$

and

$$\text{var}(X) = \frac{pq + (1-p)(1-q)}{\big(p(1-q) + q(1-p)\big)^2}.$$

Exercise:

- Two coins are simultaneously tossed until one of them comes up a head and the other a tail. The first coin comes up a head with probability p and the second with probability q. All tosses are assumed independent.
  - Find the PMF, the expected value, and the variance of the number of tosses.
  - What is the probability that the last toss of the first coin is a head?

The probability that the last toss of the first coin is a head is

$$\mathbf{P}\big(HT \mid \{HT, TH\}\big) = \frac{p(1-q)}{p(1-q) + (1-q)p}.$$

Joint PMF

$$p_{X,Y}(x,y) = \mathrm{P}(X = x \text{ and } Y = y)$$

y \ x	1	2	3	4
4	1/20	2/20	2/20	
3	2/20	4/20	1/20	2/20
2		1/20	3/20	1/20
1		1/20		

$$p_X(x) = \textstyle\sum_y p_{X,Y}(x,y)$$

$$p_{X|Y}(x|y)$$
$$= \mathrm{P}(X = x | Y = y)$$
$$= \frac{p_{X,Y}(x,y)}{p_Y(x)}$$

$$\textstyle\sum_x \sum_y p_{X,Y}(x,y) = 1$$

$$\textstyle\sum_x p_{X|Y}(x|y) = 1$$

Joint PMF: Independence

$$p_{X,Y,Z}(x,y,z) = p_X(x)p_{Y|X}(y|x)p_{Z|X,Y}(z|x,y)$$

- Random variables *X*, *Y* and *Z* are independent if (for all *x*, *y* and *z*):

$$\boxed{p_{X,Y,Z}(x,y,z) = p_X(x) \cdot p_Y(y) \cdot p_Z(z)}$$

Example:

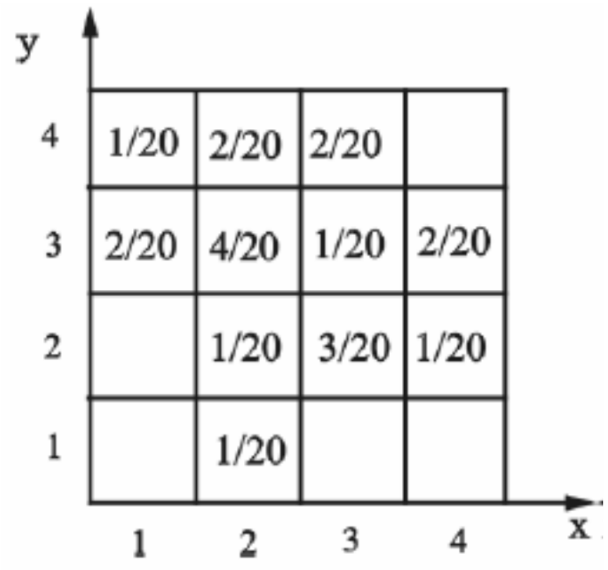
- Is X and Y Independent? **NO**

Check if $P_{X,Y}(x, y) = P_X(x).P_Y(y)$

$$P_{X,Y}(2,3) = \frac{4}{20}$$

$$P_X(2) = \frac{8}{20} \quad P_Y(3) = \frac{9}{20}$$

$$P_{X,Y}(2,3) \neq P_X(2).P_Y(3)$$



Example:

- Is X and Y Independent?
- What if we condition on $X \leq 2$ and $Y \geq 3$? **YES**

Check for all (x,y)

$$P_{X,Y}(x,y|X \leq 2, Y \geq 3) = P_X(x|X \leq 2, Y \geq 3).\, P_Y(y|X \leq 2, Y \geq 3)$$

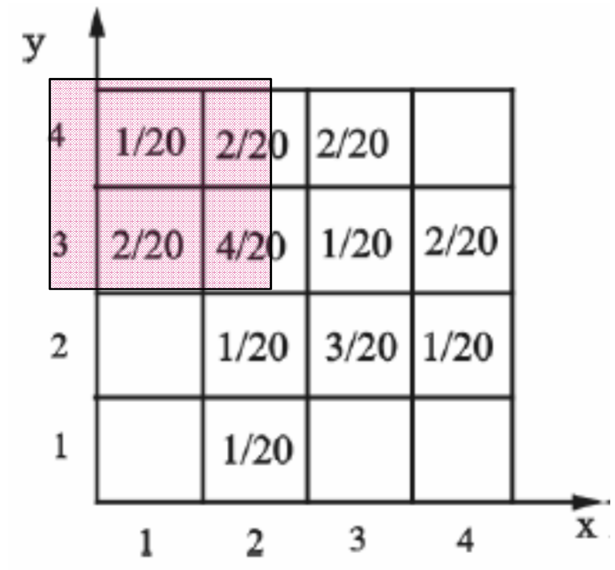
For example:

$P_{X,Y}(X,Y)\ (2,3\,|\,X\leq 2,\ Y\geq 3)=\frac{4/20}{9/20} = \frac{4}{9}$

$P_X(2\,|\,X\leq 2, Y\geq 3)=\frac{6/20}{9/20} = \frac{2}{3}$

$P_Y\ (3|X\leq 2, Y\geq 3)=\frac{6/20}{9/20}$

Check for all (x,y) to see that X and Y are independent condition on $X\leq 2$ and $Y\geq 3$



Expectations

$$\mathbf{E}[X] = \sum_x x \cdot p_X(x)$$

$$\mathbf{E}[g(X,Y)] = \sum_x \sum_y g(x,y) \cdot p_{X,Y}(x,y)$$

- In general: $\mathbf{E}[g(X,Y)] \neq g(\mathbf{E}[X], \mathbf{E}[Y])$
- $\mathbf{E}[\alpha X + \beta)] = \alpha\mathbf{E}[X] + \beta$
- $\mathbf{E}[X + Y + Z] = \mathbf{E}[X] + \mathbf{E}[Y] + \mathbf{E}[Z]$
- If X and Y are independent:
  - $\mathbf{E}[X \cdot Y] = \mathbf{E}[X] \cdot \mathbf{E}[Y]$
  - $\mathbf{E}[g(X) \cdot h(Y)] = \mathbf{E}[g(X)] \cdot \mathbf{E}[h(Y)]$

Variance

- $\text{var}(aX) = a^2\text{var}(X)$
- $\text{var}(X + a) = \text{var}(X)$
- Let $Z = X + Y$. If X and Y independent:

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$$

- Examples:
  - If $X = Y$, $\text{var}(X + Y) = 4\text{var}(X)$
  - If $X = -Y$, $\text{var}(X + Y) = 0$
  - If X, Y indep., and $Z = X - 3Y$, $\text{var}(Z) = \text{var}(X) + 9\text{var}(Y)$

Binomial Mean and Variance

- X= # of successes in n independent trials
  - Probability of success: p

$$\mathbf{E}[X] = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

- $X_i = \begin{cases} 1, & \text{if success in trial } i, \\ 0, & \text{otherwise} \end{cases}$

$\mathbf{E}[X_i] = p$

- $\text{var}(X_i) = p - p^2$

$\mathbf{E}[X] = np$

- $\text{var}(X) = np(1-p)$

The Hat Problem

- n people throw their hats in a box and then pick one at random.
  - X: number of people who get their own hat
  - Find **E**[X]

$$X_i = \begin{cases} 1, & \text{if } i \text{ selects own hat,} \\ 0, & \text{otherwise.} \end{cases}$$

$$X = X_1 + X_2 + \cdots + X_n$$

$$\mathrm{P}(X_i = 1) = \frac{1}{n}$$

$$\mathbf{E}[X_i] = \frac{1}{n}$$

Are the X_i independent? *No*

$$\mathbf{E}[X] = n(\tfrac{1}{n}) = 1$$

Variance in the Hat problem

$$\mathrm{var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \mathbf{E}[X^2] - 1$$

$$X^2 = \sum_i X_i^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} X_i X_j$$

$$\mathbf{E}[X_i^2] = \frac{1}{n}$$

$$\mathbf{P}(X_1 X_2 = 1)$$
$$= \mathbf{P}(X_1 = 1)\cdot\mathbf{P}(X_2 = 1 | X_1 = 1) = (\tfrac{1}{n})(\tfrac{1}{n-1})$$

$$\mathbf{E}[X^2] = n\frac{1}{n} + n(n-1)(\tfrac{1}{n})(\tfrac{1}{n-1}) = 2$$

$$\mathrm{var}(X) = 1$$

Example:

- Chuck will go shopping for probability books for K hours. Here, K is a random variable and is equally likely to be 1, 2, 3, or 4. The number of books N that he buys is random and depends on how long he shops. We are told that

$$p_{N|K}(n \mid k) = \frac{1}{k}, \quad \text{for } n = 1, \ldots, k$$

  - Find the joint PMF of K and N.

$$P_K(k) = \begin{cases} \dfrac{1}{4} & k = 1, 2, 3, 4 \\ 0 & k > 4 \end{cases}$$

$$P_{N,K}(n,k) = P_{N|K}(n \mid k) . P_K(k)$$

$$P_{N,K}(n,k) = \frac{1}{4k}, \quad \text{for } n = 1, \ldots, k \text{ and } k = 1, 2, 3, 4$$

Example:

- Chuck will go shopping for probability books for K hours. Here, K is a random variable and is equally likely to be 1, 2, 3, or 4. The number of books N that he buys is random and depends on how long he shops. We are told that

$$p_{N|K}(n|k) = \frac{1}{k}, \quad \text{for } n = 1,\ldots,k$$

  - Find the marginal PMF of N.

$$P_N(n) = \sum_k P_{N,K}(n,k), \quad \text{for } n = 1,\ldots,k$$

$$P_N(1) = \sum_{k=1}^{4} P_{N,K}(1,k) = \sum_{k=1}^{4} \frac{1}{4k} = \frac{1}{4} + \frac{1}{8} + \frac{1}{12} + \frac{1}{16} = \frac{25}{48}$$

$$P_N(2) = \sum_{k=2}^{4} P_{N,K}(1,k) = \sum_{k=2}^{4} \frac{1}{4k} = \frac{1}{8} + \frac{1}{12} + \frac{1}{16} = \frac{13}{48}$$

$$P_N(3) = \sum_{k=3}^{4} P_{N,K}(1,k) = \sum_{k=3}^{4} \frac{1}{4k} = \frac{1}{12} + \frac{1}{16} = \frac{7}{48} \text{ and } P_N(4) = \sum_{k=4}^{4} P_{N,K}(1,k) = \frac{1}{4k} = \frac{1}{16}$$

Example:

- Chuck will go shopping for probability books for K hours. Here, K is a random variable and is equally likely to be 1, 2, 3, or 4. The number of books N that he buys is random and depends on how long he shops. We are told that

$$p_{N|K}(n \mid k) = \frac{1}{k}, \quad \text{for } n = 1, \ldots, k$$

  - Find the conditional PMF of K given that N =2.

$$p_K(k \mid N = 2) = \frac{P_{N,K}(k, 2)}{P_N(2)}$$

$$= \frac{1/4k}{13/48} = \frac{12}{13k} \quad \text{for } k = 2, 3, 4$$

Example:

- Chuck will go shopping for probability books for K hours. Here, K is a random variable and is equally likely to be 1, 2, 3, or 4. The number of books N that he buys is random and depends on how long he shops. We are told that

$$p_{N|K}(n \mid k) = \frac{1}{k}, \quad \text{for } n = 1, \ldots, k$$

We are now told that he bought at least 2 but no more than 3 books. Find the conditional mean and variance of K, given this piece of information.

$$E(K \mid 2 \le n \le 3) = \sum_k k.P_K(k \mid 2 \le n \le 3) \text{ and } P_K(k \mid 2 \le n \le 3) = \frac{P_{N,K}(2,k) + P_{N,K}(3,k)}{P_N(2) + P_N(3)}$$

$$P_K(1 \mid 2 \le n \le 3) = 0 \quad \text{and} \quad P_K(2 \mid 2 \le n \le 3) = \frac{1/8}{5/12} = \frac{3}{10}$$

$$P_K(3 \mid 2 \le n \le 3) = \frac{1/6}{5/12} = \frac{2}{5} \quad \text{and} \quad P_K(4 \mid 2 \le n \le 3) = \frac{1/8}{5/12} = \frac{3}{10}$$

$$E(K \mid 2 \le n \le 3) = \frac{2*3}{10} + \frac{3*2}{5} + \frac{4*3}{10} = 3$$

Example:

- Chuck will go shopping for probability books for K hours. Here, K is a random variable and is equally likely to be 1, 2, 3, or 4. The number of books N that he buys is random and depends on how long he shops. We are told that

$$p_{N|K}(n \mid k) = \frac{1}{k}, \quad \text{for } n = 1, \ldots, k$$

  - We are now told that he bought at least 2 but no more than 3 books. Find the conditional mean and variance of K, given this piece of information.

$$E(K \mid 2 \le n \le 3) = \frac{2*3}{10} + \frac{3*2}{5} + \frac{4*3}{10} = 3$$

$$E(K^2 \mid 2 \le n \le 3) = \frac{4*3}{10} + \frac{9*2}{5} + \frac{16*3}{10} = \frac{58}{5}$$

$$Var(K \mid 2 \le n \le 3) = E(K^2 \mid 2 \le n \le 3) - [E(K \mid 2 \le n \le 3)]^2$$

$$= \frac{58}{5} - 9 = \frac{13}{5}$$