Probability and Statistics: Lesson 3
Random Variables: Discrete

Morgan McCarty

10 July 2023

1 Lesson 1 Addendum

1.1 Counting

1.1.1 Defiinitions

- Combination: nCk or $\binom{n}{k}$ is the number of ways to choose k objects from a set of n objects.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{nPk}{k!} \tag{1}$$

- Permutation: nPk is the number of ways to arrange k objects from a set of n objects.

$$nPk = \frac{n!}{(n-k)!} = k!\binom{n}{k} \tag{2}$$

1.1.2 Examples

- Given the set $S = \{A, B, C\}$ how many ways can we choose 3 letters from S? How many ways can we arrange those letters?
  - $3C3 = 1$
  - $3P3 = 6$
- There is only one way to choose 3 letters from S because there are only 3 letters in S. However, there are 6 ways to arrange those letters.

2 Random Variable: Discrete

2.1 Definition

- Random Variable: a function that maps the sample space to the real numbers. $X : S \to \mathbb{R}$
- Probability Density Functon or Probability Distributon: $p_X(x) = P(X = x)$ or $p_X(k) = P(X = k)$. Often written as a table.

- Cumulative Distribution Function: $F_X(x) = P(X \leq x)$ or $F_X(k) = P(X \leq k)$. Often written as a table.

$$F_X(x) = \sum_{k \leq x} p_X(k) \tag{3}$$

- Expected Value: $E[X] = \mu_X = \sum_k k \cdot p_X(k)$
- Range: the set of all possible values of X. $Ran(X) = \{x \in \mathbb{R} | p_X(x) > 0\}$.
- Variance: $Var(X) = \sigma_X^2 = E[(X - \mu_X)^2] = \sum_k (k - \mu_X)^2 \cdot p_X(k)$
  The useful equation $Var(X) = E(X^2) - \mu^2$ is derived in the proof below.
  Proof of $Var(X) = E(X^2) - \mu^2$:

$$\begin{aligned} Var(X) = E((X-\mu)^2) &= E(X^2 - 2X\mu + \mu^2) \\ &= E(X^2) - 2\mu E(X) + E(\mu^2) \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

  The Variance is the expected value of the squared difference between X and the expected value of X.
  In layman's terms, the variance is the average squared distance from the mean.

- Standard Deviation: $\sigma_X = \sqrt{Var(X)}$
  The standard deviation is the square root of the variance or the average distance from the mean.

2.2 Examples

1. Example: 2 Spinning Wheels

   - Given 2 3-way spinning wheels, find the probability distribution of the sum of the two wheels: X.

$$S = \begin{Bmatrix} (1,1), & (1,2), & (1,3), \\ (2,1), & (2,2), & (2,3), \\ (3,1), & (3,2), & (3,3) \end{Bmatrix}$$

   - The probability distribution of X is:

k	2	3	4	5	6
$p_X(k)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

   - The cumulative distribution of X is:

k	2	3	4	5	6
$F_X(k)$	$\frac{1}{9}$	$\frac{3}{9}$	$\frac{6}{9}$	$\frac{8}{9}$	$\frac{9}{9}$

- The expected value of X is:

$$\begin{aligned}
E(X) &= \mu_X \\
&= \sum_k k \cdot p_X(k) \\
&= (2 \cdot \frac{1}{9}) + (3 \cdot \frac{2}{9}) + (4 \cdot \frac{3}{9}) + (5 \cdot \frac{2}{9}) + (6 \cdot \frac{1}{9}) \\
&= \frac{(2+6+12+10+6)}{9} \\
&= \frac{36}{9} \\
&= 4
\end{aligned}$$

- The variance of X is:

$$\begin{aligned}
Var(X) &= \sigma_X^2 \\
&= E(X^2) - \mu^2 \\
&= 2^2 \cdot \frac{1}{9} + 3^2 \cdot \frac{2}{9} + 4^2 \cdot \frac{3}{9} + 5^2 \cdot \frac{2}{9} + 6^2 \cdot \frac{1}{9} - 4^2 \\
&= \frac{(4+18+48+50+36)}{9} - 16 \\
&= \frac{156}{9} - 16 \\
&= \frac{12}{9} \\
&= \frac{4}{3}
\end{aligned}$$

- The standard deviation of X is:

$$\begin{aligned}
\sigma_X &= \sqrt{Var(X)} \\
&= \sqrt{\frac{4}{3}} \\
&= \frac{2}{\sqrt{3}}
\end{aligned}$$

2. Example: 3 Fair Dice

- Given 3 fair dice, X is the largest number rolled.
- While usually, the probability density function is easier to find than the cumulative distribution function, in this case, the opposite is true.
- Additional both can be written as equations rather than tables.
- The cumulative distribution function of X is:

$$F_X(k) = \frac{k^3}{6^3}$$

- The probability density function of X is:

$$p_X(k) = F_X(k) - F_X(k-1)$$

2.3 Bernouli Variable (Distribution)

- Bernouli Variable: a random variable that can only take on two values, 0 or 1.
- Bernouli Distribution: the probability distribution of a Bernouli Variable.

$$X \sim Bernouli(p)$$

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

$$F_X(k) = \begin{cases} 0 & k < 0 \\ 1 - p & 0 \leq k < 1 \\ 1 & k \geq 1 \end{cases}$$

$$\begin{aligned} E(X) &= 0 \cdot (1 - p) + 1 \cdot p \\ &= p \end{aligned}$$

$$\begin{aligned} Var(X) &= E(X^2) - \mu^2 \\ &= 0^2 \cdot (1 - p) + 1^2 \cdot p - p^2 \\ &= p - p^2 \\ &= p(1 - p) \end{aligned}$$