ISyE 3770, Spring 2024
Statistics and Applications

Introduction to Random Variables

Instructor: Jie Wang
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

jwang3163@gatech.edu
Office: ISyE Main 445

Discrete Distribution

Section 2.1 **Random variable** of the discrete type

$$\text{Outcomes of experiment} \begin{cases} \text{Numerical} \\ \text{e.g. Rolling a 6-sided die} \\ \text{Not numerical} \\ \text{e.g. Flipping a coin} \end{cases}$$

For the latter case, we can define a function X to associate the outcomes with numerical values.

Example 1: Rolling a die: $S = \{1,2,3,4,5,6\}$.

$X(i) = i, \quad i = 1,2,3,4,5,6.$

Flipping a coin: $S = \{H, T\}$.

$X(H) \triangleq 0, X(T) \triangleq 1.$

Definition [Random Variable (RV)]

- Given a random experiment with sample space S, a function $X: S \to B \subseteq \mathbb{R}$ that assign one and only one real number $X(s) = x$ for each $s \in S$ is called **random variable**.
- In other words, A ***random variable*** is a function from a sample space S into the real numbers.

Definition [Discrete Random Variable]

The range of X is the set

$$B = \{ x \mid X(s) = x, s \in S \}.$$

A RV is called **discrete** if its range B is finite or countable.

Given an experiment with sample space

$$S = \{s_1, \dots, s_n\}$$

with a probability function P_r on S, and we define a random variable X with range $B = \{x_1, \dots, x_m\}$, we can define a probability function P on B in the following way:

$$P(X = x_i) \triangleq P(\{X = x_i\}) = P_r(\{s_j \mid X(s_j) = x_i, s_j \in S\})$$

$$P(X \in A) \triangleq P(\{X \in A\}) = P_r(\{s_j \mid X(s_j) \in A, s_j \in S\})$$

Note that $A \subseteq B$

Notation Remark: Random variables will always be denoted with uppercase letters. The numerical values of RV will be denoted by the corresponding lowercase letters

$$X \rightarrow \text{a RV}, \qquad x \rightarrow \text{the numerical value of a RV.}$$

Thus, the random variable X can take the value x.

Definition [probability mass function (pmf)]

Suppose that $X : S \to B \subseteq \mathbb{R}$ is a discrete random variable. Then a function $f(x) : B \to [0, 1]$ is called a pmf, if

- $f(x) > 0, x \in B$;
- $\sum_{x\in B} f(x) = 1$;
- $P(X \in A) = \sum_{x\in A} f(x)$, where $A \subseteq B$.

- We often extend the definition domain of $f(x)$ from B to $\mathbb{R}$ and let $f(x)=0$ for $x \notin B$.
- B is the **range** of X and is also called the **support** of $f(x)$.
- From now on, we consider pmf $f(x)$: $\mathbb{R} \to [0,1]$.

Definition [Cumulative distribution function (cdf)]

The cumulative distribution function or cdf of a random variable X, denoted by $F(x)$, is defined by

$$F(x) = P(X \le x) \triangleq P(\{s \mid X(s) \le x, s \in S\}), \quad x \in (-\infty, \infty).$$

- **Remark**: cdf of X is also called the distribution function of X

Definition [uniform distribution]
When a pmf is constant over the support.

Example 2: Rolling a die. $S = \{1,2,3,4,5,6\} \to B = \{1,2,3,4,5,6\}$

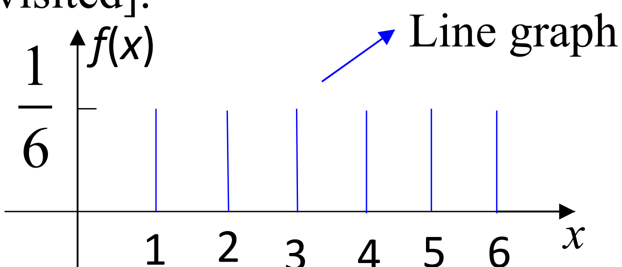
- Define a RV $X(s) = s$ for $\forall s \in S$.

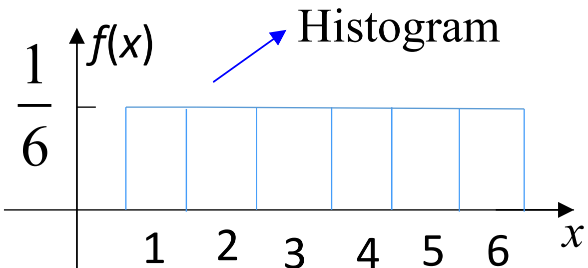
- pmf $f(x) = \begin{cases} 1/6, & \text{if } x \in B, \\ 0, & \text{if } x \notin B. \end{cases}$

- cdf $F(x) = P(X \leq x) = \begin{cases} 0, & \text{if } x < 1, \\ k/6, & \text{if } k \leq x < k+1, k = 1,2,3,4,5, \\ 1, & \text{if } x \geq 6 \end{cases}$

Definition [line graph]
A line graph of the pmf $f(x)$ of the random variable X is a graph having a vertical line segment drawn from $(x, 0)$ to $(x, f(x))$ at each $x \in S$.

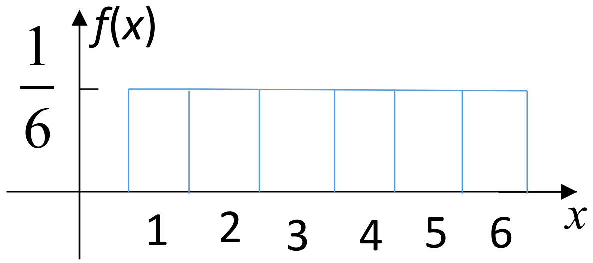
Example 2 [Revisited]:





Definition [probability histogram]

If a RV X assumes only integer values, a probability histogram of pmf $f(x)$ is a graphical representation that has a rectangle of height $f(x)$ and a base of length 1, centered at x for each $x \in S$.



Probability Histogram

Section 2.2 Mathematical expectation

We will learn many probability distributions, it's important to introduce concepts in summarizing their key characteristics.

- Expectation
- Variance
- …

➢ **Motivation Example.**

A man proposes a game: let the other player throw a die and the player receives payment as follows:

$$\boxed{\begin{aligned} A &= \{1,2,3\} \to 1 \text{ dollar} \\ B &= \{4,5\} \to 2 \text{ dollars} \\ C &= \{6\} \to 3 \text{ dollars} \end{aligned}}$$

$$\boxed{X : \ S \triangleq A \bigcup B \bigcup C \to X(S) = \{1,2,3\}.}$$

Now let X be a RV to represent the payment, the pmf of X is:

$$\boxed{\begin{aligned} f(x) &= \frac{4-x}{6}, \quad x = 1,2,3 \\ f &: X(S) \to \{1/6, 1/3, 1/2\} \end{aligned}}$$

- The man charge the player 2 dollars for each play. Can the man make profit if the game is repeated endlessly?

Solution. Payment of $\begin{bmatrix}1\\2\\3\end{bmatrix}$ occurs $\begin{bmatrix}3/6\\2/6\\1/6\end{bmatrix}$ of the times.

Longrun average value of X

The average payment is $\boxed{1\cdot 3/6 + 2\cdot 2/6 + 3\cdot 1/6 = 5/3.}\triangleq \boldsymbol{E(X)}$
So the man can earn $2 - 5/3 = 1/3$ per play on average.

More generally, we are interested in long run average value of a function of X, say $g(X)$.

Definition [Expectation]

Assume X is a discrete RV with range space $X(S)$ and $f(x)$ is its pmf. If $\sum_{x\in X(S)} g(x)f(x)$ exists, then it is called the **expectation** or the **expected value** of $g(X)$, denoted as $\mathbb{E}[g(X)]$. That is,

$$\mathbb{E}[g(X)] = \sum_{x\in X(S)} g(x)f(x).$$

Example 1

Let X be a RV with $X(S) = \{-1, 0, 1\}$ and its pmf is $f(x) = 1/3$ for any $x \in X(S)$. What is $\mathbb{E}[X^2]$?

Solution. $\mathbb{E}[X^2] = \sum_{x \in X(S)} x^2 f(x) = (-1)^2 \cdot 1/3 + 0^2 \cdot 1/3 + 1^2 \cdot 1/3 = 2/3$.

➢ Properties of mathematical expectation

Theorem 2.2-1

Consider a RV $X : S \to X(S)$ and its pmf $f : X(S) \to [0, 1]$. When the mathematical expectation exists, it satisfies the following properties:

- If c is a constant, then $\mathbb{E}[c] = c$.
- If c is a constant, and g is a function,

$$\mathbb{E}[cg(X)] = c\mathbb{E}[g(X)].$$

- If c_1, c_2 are constants, and g_1, g_2 are functions,

$$\mathbb{E}[c_1 g_1(X) + c_2 g_2(X)] = c_1 \mathbb{E}[g_1(X)] + c_2 \mathbb{E}[g_2(X)].$$

Example 2

Let $g(x) = (x - b)^2$, where b is a constant to be chosen. Suppose $\mathbb{E}[(X - b)^2]$ exists. Find the value of b such that $\mathbb{E}[(X - b)^2]$ is minimal.

Solution. Notice that

$$\begin{aligned} h(b) \triangleq \mathbb{E}[(X - b)^2] &= \mathbb{E}[X^2 - 2b \cdot X + b^2] \\ &= \mathbb{E}[X^2] - 2b \cdot \mathbb{E}[X] + b^2. \end{aligned}$$

Besides,

$$\frac{\partial h(b)}{\partial b} = -2\mathbb{E}[X] + 2b, \quad \frac{\partial^2 h(b)}{\partial b^2} = 2 > 0.$$

Therefore, when $\frac{\partial h(b)}{\partial b} = 0$, $\mathbb{E}[(X - b)^2]$ is minimal. Then $b = \mathbb{E}[X]$.

Mean is the Minimum Mean Squared Error (MMSE) estimator.

Section 2.3 Special mathematical expectation

➢ Mean of RV: The expectation of X is also called the mean of X.

$$\mathbb{E}[X] = \sum_{x \in X(S)} x f(x) = \sum_{i=1}^{k} u_i f(u_i)$$

Assume that
$X(S) = \{u_1, \dots, u_k\}$

➢ Mechanic Interpretation:

- u_i: the distance of the i-th point from the origin.
- $f(u_i)$: the weight of the i-th point.
- $u_i f(u_i)$: a moment having a moment arm of length u_i.
- $\mathbb{E}[X]$: the 1st order moment about the system; the centroid.

Why $\mathbb{E}[X]$ is the centroid?

If we choose $\mathbb{E}[X]$ as the new origin, then we compute the 1st order moment again:

$$\mathbb{E}\Big[X - E[X]\Big] = 0.$$

Hence, the 1st order moment about $\mathbb{E}[X]$ is zero, i.e., $\mathbb{E}[X]$ is centroid.

Variance of RV

$$
\begin{aligned}
\mathbb{V}\mathrm{ar}[X] \triangleq \boxed{\mathbb{E}\Big[(X-\mathbb{E}[X])^2\Big]} &= \sum_{x \in X(S)} (x-\mathbb{E}[X])^2 f(x) \\
&= \mathbb{E}\Big[X^2 - 2X\mathbb{E}[X] + (\mathbb{E}[X])^2\Big] \\
&= \mathbb{E}[X^2] - 2\mathbb{E}[X\mathbb{E}[X]] + (\mathbb{E}[X])^2 \\
&= \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + (\mathbb{E}[X])^2 \\
&= \boxed{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}
\end{aligned}
$$

- The positive square root of the variance of the RV X is called the **standard deviation**, denoted as δ_X.

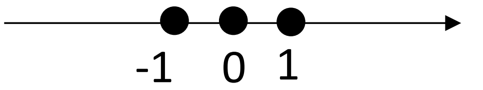
Example 1:

Let X equal to the number of spots after a 6-sided die is rolled. The probability model is

$$f(x) = P(X = i) = 1/6, \quad i = 1, 2, \ldots, 6.$$

Mean of X: $\mathbb{E}[X] = 1/6 * (1 + 2 + \cdots + 6) = 7/2$.
Variance of X: $\mathbb{V}\mathrm{ar}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$

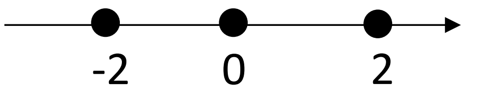
Example 2 [Interpretational standard deviation]



Let X have pmf $f(x) = 1/3$ for $x = -1, 0, 1$, then

$$\mathbb{E}[X] = 0, \quad \mathbb{V}\mathrm{ar}[X] = 2/3, \quad \delta_X = \sqrt{2/3}.$$

Let Y have pmf $f(x) = 1/3$ for $x = -2, 0, 2$, then



$$\mathbb{E}[X] = 0, \quad \mathbb{V}\mathrm{ar}[X] = 8/3, \quad \delta_X = 2\sqrt{2/3}.$$

Standard deviation is a measure of the dispersion or spread of the points belonging to the range space of RV.

- Properties of variance

Let X be a RV, then

$$\mathbb{V}\mathrm{ar}[c] = 0, \quad \mathbb{V}\mathrm{ar}[cX] = c^2\mathbb{V}\mathrm{ar}[X].$$

r-th moment of the distribution

- Let r be a positive integer. If $\mathbb{E}[X^r] = \sum_{x \in X(S)} x^r f(x)$ exists and is finite, then it is called the **r-th moment of the distribution about the origin**.

- In addition, $\mathbb{E}[(X-b)^r] = \sum_{x \in X(S)} (x-b)^r f(x)$ is called the **r-th moment of the distribution about b**.

- $\mathbb{E}[(X)_r] \triangleq \mathbb{E}[X(X-1)\cdots(X-r+1)]$ is called the **r-th factorial moment**.

- Mean
- Variance
- Standard deviation
- moments

→ Characteristics of distribution of probability

We now define a function that will help us generate the moments of a distribution:

Definition [Moment generating function (mgf)]

Let X be a discrete RV with range space $X(S)$. If there exists $h > 0$ such that $\mathbb{E}[e^{tX}] = \sum_{x \in X(S)} e^{tx} f(x)$ exists and is finite for $-h < t < h$, then the function defined by $M(t) = \mathbb{E}[e^{tX}]$ is called the **moment-generating function** of X.

- Properties of mgf:
  I. $M(0) = 1$.
  II. If two RVs have the same mgf, they must have the same distribution of probability.

Example 3:

Suppose X has the mgf

$$M(t) = e^t \cdot \frac{3}{6} + e^{2t} \cdot \frac{2}{6} + e^{3t} \cdot \frac{1}{6}, \quad -\infty < t < \infty,$$

then the support of the pmf $f(x)$ of X is $S = \{1, 2, 3\}$, and the associated pmf $f(x) = \frac{4-x}{6}, x = 1, 2, 3$.

$$M'(t) = \sum_{x \in X(S)} x e^{tx} f(x)$$

$$M''(t) = \sum_{x \in X(S)} x^2 e^{tx} f(x)$$

$$M^{(r)}(t) = \sum_{x \in X(S)} x^r e^{tx} f(x)$$

Noted question:

- Is $M(t)$ differentiable? in 1st, 2nd,…order
- Interchange of the differentiation and summation.

Putting $t = 0$, we find $M'(0) = \mathbb{E}[X], M''(0) = \mathbb{E}[X^2], M^{(r)}(0) = \mathbb{E}[X^r]$.

Remark: The moments can be computed by differentiating $M(t)$!

Example 4:

Suppose X has the geometric distribution, that is, the pmf of X is

$$f(x) = q^{x-1}p, \quad x = 1, 2, \ldots, n, \ldots, \quad p \triangleq 1 - q.$$

Then the mgf of X is

$$M(t) = \mathbb{E}[e^{tX}] = \sum_{x=1}^{\infty} e^{tx} \cdot q^{x-1}p = \frac{p}{q}\sum_{x=1}^{\infty}(qe^t)^x$$

$$= \frac{p}{q}\left[(qe^t) + (qe^t)^2 + \cdots\right]$$

$$= \frac{p}{q}\frac{qe^t}{1-qe^t} = \frac{pe^t}{1-qe^t}, \quad \text{provided that } qe^t < 1 \Longleftrightarrow t < -\ln q.$$

To find the mean and variance of X,

$$\boxed{\begin{aligned} M'(t) &= \frac{pe^t}{1-qe^t} - \frac{(pe^t)(-qe^t)}{(1-qe^t)^2} = \frac{pe^t}{(1-qe^t)^2} \\ M''(t) &= \frac{pe^t(1+qe^t)}{(1-qe^t)^3} \end{aligned}}$$

$$M'(0) = \mathbb{E}[X] = \frac{p}{(1-q)^2} = \frac{1}{p}$$

$$M''(0) = \mathbb{E}[X^2] = \frac{1+q}{p^2}$$

$$\mathbb{V}\mathrm{ar}[X] = \mathbb{E}[X^2] - \left(\mathbb{E}[X]\right)^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$