Topic 7
Random Variables and Distribution Functions

Distribution Functions
Outline

Definition of a Random Variable

Distribution Functions
  Discrete Random Variables
  Continuous Random Variables

Properties of Distribution Functions
**Introduction**

<table>
<thead>
<tr>
<th>statistics</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>universe of information</td>
<td>sample space - $\Omega$</td>
</tr>
<tr>
<td>$\Downarrow$</td>
<td>and probability - $P$</td>
</tr>
<tr>
<td>ask a question and collect data</td>
<td>$\Downarrow$ define a random variable $X$</td>
</tr>
<tr>
<td>$\Downarrow$</td>
<td>$\Downarrow$ organize into the cumulative distribution function</td>
</tr>
<tr>
<td>organize into the empirical cumulative distribution function</td>
<td>$\Downarrow$ organize into the cumulative distribution function</td>
</tr>
<tr>
<td>$\Downarrow$</td>
<td>$\Downarrow$ compute distributional means and variances</td>
</tr>
<tr>
<td>compute sample means and variances</td>
<td>$\Downarrow$ compute distributional means and variances</td>
</tr>
</tbody>
</table>
Definition of a Random Variable

A random variable is a real valued function from the probability space.

\[ X : \Omega \to \mathbb{R}. \]

Typically, we shall use capital letters near the end of the alphabet, e.g., \( X, Y, Z \) for random variables. The range of a random variable is called the state space.

Exercise. Give some random variables on the following probability spaces, \( \Omega \).

1. Roll a die 3 times and consider the sample space

\[ \Omega = \{(i, j, k); i, j, k = 1, 2, 3, 4, 5, 6\} . \]

2. Flip a coin 10 times and consider the sample space \( \Omega \), the set of 10-tuples of heads and tails.

We can create new random variables via composition of functions:

\[ \omega \mapsto X(\omega) \mapsto g(X(\omega)) \]

Thus, if \( X \) is a random variable, then so are

\[ X^2, \quad \exp \alpha X, \quad \sqrt{X^2 + 1}, \quad \tan^2 X, \quad [X] \]
Definition of a Random Variable

Distribution Functions

A (cumulative) distribution function of a random variable $X$ is defined by

$$F_X(x) = P\{\omega \in \Omega; X(\omega) \leq x\} = P\{X \leq x\}.$$

For the complement of $\{X \leq x\}$, we have the survival function

$$\bar{F}_X(x) = P\{X > x\} = 1 - P\{X \leq x\} = 1 - F_X(x).$$

Choose $a < b$, then the event $\{X \leq a\} \subset \{X \leq b\}$. Their set theoretic difference

$$\{X \leq b\} \setminus \{X \leq a\} = \{a < X \leq b\}.$$

Consequently, by the difference rule for probabilities,

$$P\{a < X \leq b\} = P(\{X \leq b\} \setminus \{X \leq a\}) = P\{X \leq b\} - P\{X \leq a\} = F_X(b) - F_X(a).$$

In particular, $F_X$ is non-decreasing.
Definition of a Random Variable

Distribution Functions

Let $X$ be the sum of the values on two fair dice,

\[
\begin{array}{cccccccccccc}
 x & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\
\end{array}
\]

Here is $F_X$ the cumulative distribution function.
Distribution Functions

Notice that the distribution function

• is constant in between the possible values for $X$,
• has a jump size at $x$ is equal to $P\{X = x\}$, and
• is right continuous.

Call $X$ a discrete random variable if its distribution function $F_X$ has these properties.

Examples.

\[
\frac{3}{36} = P\{X = 4\} = F_X(4) - F_X(4-) = \frac{6}{36} - \frac{3}{36}
\]

\[
P\{4 < X \leq 7\} = F_X(7) - F_X(4) = \frac{21}{36} - \frac{6}{36} = \frac{15}{36} = \frac{5}{12}.
\]

\[
P\{4 \leq X \leq 7\} = F_X(7) - F_X(4-) = \frac{21}{36} - \frac{3}{36} = \frac{18}{36} = \frac{1}{2}.
\]
Distribution Functions

Exercise.

1. Flip a fair coins 3 times. Let $X$ be the number of heads. Under equally likely outcomes, find

$$P\{X = x\} \quad \text{for} \quad x = 0, 1, 2, \text{and} \ 3.$$ 

and use this to sketch a graph of the distribution function $F_X$.

2. Deal 5 cards out of a deck of 52. Let $X$ be the number of ♦. Under equally likely outcomes, use the choose function in R to determine

$$P\{X = x\} \quad \text{for} \quad x = 0, 1, 2, 3, 4, \text{and} \ 5.$$ 

and use this to sketch a graph of the distribution function $F_X$. 
Distribution Functions

For a dart board with radius 1, assume that the dart lands randomly uniformly. Let $X$ be the distance from the center. For $x \in [0, 1],$

$$F_X(x) = P\{X \leq x\} = \frac{\text{area inside circle of radius } x}{\text{area of circle}} = \frac{\pi x^2}{\pi 1^2} = x^2.$$  

Thus, we have the distribution function

$$F_X(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ x^2 & \text{if } 0 < x \leq 1, \\ 1 & \text{if } x > 1. \end{cases}$$
Exercise.

1. Find the probability that the dart no more than 1/2 unit from the center.
2. Find the probability that the dart lands further 1/3 unit but no more than 2/3 unit from the center.
3. Find the median, $x_{1/2}$ so that $P\{X \leq x_{1/2}\} = 1/2$.

Definition. $X$ is continuous random variable if it has a cumulative distribution function $F_X$ that is differentiable.
Properties of Distribution Functions

A distribution function $F_X$ has the property that it is right continuous, starts at 0, ends at 1, and does not decrease with increasing values of $x$.

In mathematical terms,

- For every $a$, $\lim_{x \to a^+} F_X(x) = F_X(a)$.
- $\lim_{x \to -\infty} F_X(x) = 0$.
- $\lim_{x \to \infty} F_X(x) = 1$.
- For every $a, b$ satisfying $a < b$, $F_X(a) \leq F_X(b)$. 