Today: Random variables

Next: ASV 3.2

Week 3:

- homework 1 graded, regrades open until October 16
- homework 3 (due Friday, October 21)
- midterm 1: Wednesday, October 19
Cumulative Distribution Function (CDF)

For any random variable $X$, define

CDF: $F_X(r) = P(X \leq r)$

Example: $X \sim \text{Bin}(3, \frac{1}{2})$
Properties of the CDF \[\text{CDF} \quad F_X(r) = P(X \leq r)\]

(1) Monotone increasing:

(2) \[
\lim_{r \to -\infty} F_X(r) = 0, \quad \lim_{r \to +\infty} F_X(r) = 1
\]

(3) The function \( F_X \) is right-continuous:

\[
\lim_{t \to r^+} F_X(t) = F_X(r)
\]

Corollary: If \( X \) is a continuous random variable, \( F_X \) is a

Example: Shoot an arrow at a circular target of radius 1 (choose point from unit disk uniformly at random)

\[
F_X(r) = \begin{cases} 
0, & r \leq 0 \\
0, & 0 < r \leq 1 \\
1, & r \geq 1 
\end{cases}
\]
Cumulative distribution function (CDF)

Summary: For any random variable X, \( F_X(r) = P(X \leq r) \)

1. Monotone increasing: \( s \leq t \Rightarrow F_X(s) \leq F_X(t) \)

2. \( \lim_{r \to -\infty} F_X(r) = 0 \), \( \lim_{r \to +\infty} F_X(r) = 1 \)

3. Right-continuous: \( \lim_{t \to r^+} F_X(t) = F_X(r) \)

Discrete random variable

Finite or countable set of values with \( t_1, t_2, \ldots, P(X = t_j) > 0 \) and \( \sum_{j} P(X = t_j) = 1 \)

Continuous random variable

For each real number \( t \), \( P(X = t) = 0 \)

Because (1) and (3) this implies that \( F_X \) is continuous

no jumps
## Densities (PDF)

Some continuous random variables have probability densities. This is the infinitesimal version of the probability mass function.

<table>
<thead>
<tr>
<th>$X$ discrete, $X \in {t_1, t_2, \ldots}$</th>
<th>$X$ continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_x(t) = P(X = t)$</td>
<td>$P(X = t) = 0$ for all $t \in \mathbb{R}$</td>
</tr>
</tbody>
</table>

probability mass function
Example: Shoot an arrow at a circular target of radius 1.

\[ X = \text{distance from center} \]

\[ F_X (r) = \begin{cases} 
0, & r \leq 0 \\
r^2, & 0 < r \leq 1 \\
1, & r \geq 1 
\end{cases} \]
Thm: If \( F_x \) is continuous and (piecewise) differentiable, then \( X \) has density

Proof: Follows from FTC

Example: Let \( X \) = random number chosen uniformly on \([0,1]\).

We have seen that in this case \( P(X \in [s,t]) = t-s, \ 0 \leq s < t \leq 1 \).

\[
F_x(r) = P(X \leq r) = \begin{cases} \end{cases}

f_x(r) = \begin{cases} \end{cases}

PDF

Example Let \( f(t) = \begin{cases} c \sqrt{1-t^2}, & 1 \leq |t| \\ 0, & \text{otherwise} \end{cases}, \quad c > 0 \)

Q: (When) Is \( f(t) \) a PDF of some random variable?

- \( f \geq 0 \)

- \( +\infty \) \[ 1 = \int_{-\infty}^{+\infty} f(t) \, dt \]

\( f \) is a PDF
Question

Your car is in a minor accident. The damage repair cost is a random number between 100 and 1500 dollars. Your insurance deductible is 500 dollars.

\[ Z = \text{your out of pocket expenses} \]

Question: The random variable \( Z \) is

(a) continuous

(b) discrete

(c) neither

(d) both
Independent random variables

A collection $X_1, X_2, \ldots, X_n$ of random variables defined on the same sample space are independent if for any $B_1, B_2, \ldots, B_n \subset \Omega$, the events

i.e.,

Special case: if $X_j$ are discrete random variables, it suffices to check the simpler condition for any real numbers $t_1, t_2, \ldots, t_n$

Example Let $X_1, X_2, \ldots, X_n$ be fair coin tosses, $H \sim 1$, $T \sim 0$

$$P(X_1 = t_1, X_2 = t_2, \ldots, X_n = t_n) = P(X_1 = t_1) \cdots P(X_n = t_n)$$
Bernoulli distribution

Experiments can have numerical observables, but sometimes you only observe whether there is success or failure. We model this with a random variable $X$ taking value 1 with probability $p$, and value 0 with probability $1-p$.

In practice, we usually repeat the experiment many times, making sure to use the same setup each trial. The previous trials do not influence the future ones.
Independent trials. Binomial distribution

Let $X_1, X_2, \ldots, X_n$ be independent $\text{Ber}(p)$ random variables.

E.g. 

$$P(X_1 = 0, X_2 = 1, X_3 = 1, X_4 = 0, X_5 = 0, X_6 = 0)$$

= 

= 

Run $n$ independent trials, each with success probability $p$.

Let $X_1, X_2, \ldots, X_n \sim \text{Ber}(p)$

Let $S_n = \# \text{ successful trials}$

What is the distribution of $S_n$?

$$P(S_n = k) = P(\{\text{exactly } k \text{ of the } n \text{ trials are successful}\})$$

= 

If 

, # of Heads in $n$ tosses