Math 180A: Introduction to Probability

Today: ASV 3.1 (Probability distributions)

ASV 3.2 (Cumulative distribution function)

Video: Prof. Todd Kemp, Fall 2019

Next: ASV 3.2

Week 3: Quiz 2 (Wednesday, Jan 20, on Lectures 3-5)
Homework 2 (due Friday, Jan 22)
Random Variables

Given a probability space \((\Omega, \mathcal{F}, P)\), a random variable is a function

\[ X : \Omega \rightarrow \mathbb{R} \]

This is a bad, old-fashioned name. Would be better to call it a random function or random measurement.

Eg. Toss a fair coin 4 times. Let \( X \) = number of tails.

Eg. Shoot an arrow at a circular target. \( Y \) = distance from center.

Eg. Your car is in a minor accident; the damage repair cost is a random number between $100 and $1500. Your insurance deductible is $500. \( Z \) = your out of pocket expenses.

In all these examples, think about what you observe. You can't really see a formula for the function. By repeating the experiment over and over, all you can learn is the probability distribution.
Probability Distribution

Given a probability space \((\Omega, \mathcal{F}, P)\) and a random variable

\[ X: \Omega \rightarrow \mathbb{R} \]

the probability distribution or law of \(X\) is a probability measure \(\mu_X\) on \(\mathbb{R}\).

\[ A \subseteq \mathbb{R} \quad \Rightarrow \quad \mu_X(A) = P(\{x \in A\}) \]

[Caution: for this to make sense, we need to have a designated set of allowed "events" in \(\mathbb{R}\); call this collection \(\mathcal{B}(\mathbb{R})\).

Then, we must have

\[ \text{For each } A \in \mathcal{B}(\mathbb{R}), \{x \in A\} \in \mathcal{F}. \]

This is a condition on \(X\); we call such functions measurable.

We will ignore these technicalities in this course; all our random variables are indeed measurable.
E.g. Toss a fair coin 4 times. Let $X = \text{number of tails.}$

$\Omega = \{(x_1, x_2, x_3, x_4) \in \{\text{H, T}\}^4\}$

$\mathbb{P} = \text{uniform on } \Omega$;

$\mathbb{P}\{(x_1, x_2, x_3, x_4)\} = \frac{1}{2^4} = \frac{1}{16}$

$\{X = 2\} = \{(T, T, H, H), \ldots\}$

$\#\{X = 2\} = \binom{4}{2} = \frac{4!}{2!2!} = \frac{3}{8}$

$\mathbb{P}(X = 2) = \frac{\binom{4}{2}}{16} = \frac{3}{8}$

$\mathbb{P}(X = k) = \frac{\binom{4}{k}}{16}$

$X = \#\text{tails in } n \text{ coin tosses,}$

$p_x(k) = \mathbb{P}(X = k) = \frac{1}{2^n} \binom{n}{k}$

$k \quad 0 \quad 1 \quad 2 \quad 3 \quad 4$

$p_x(k) \quad \frac{1}{16} \quad \frac{1}{4} \quad \frac{3}{8} \quad \frac{1}{4} \quad \frac{1}{16}$

$\text{Binomial } \text{Bin}(n, \frac{1}{2})$
Eg. Shoot an arrow at a circular target of radius 1.

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

Last lecture, you calculated that:

- If \( r \leq 0 \), \( \Pr(Y \leq r) = 0 \)
- If \( r > 1 \), \( \Pr(Y \leq r) = 1 \)
- If \( r \leq 1 \), \( \Pr(Y \leq r) = \frac{\text{Area}(\text{IDr})}{\text{Area}(\text{ID})} = r^2 \)

\[ \text{as } \varepsilon \to 0, \text{ this } \to 0. \]

What can we say about \( \Pr(Y = 0.4) \)?

\( (0.3, 0.4] \)

\[ \mu_Y(-\infty, 0.4] = \mu_Y(-\infty, 0.3] + \mu_Y(0.3, 0.4] \]

\[ \Pr(Y \in (0.4 - \varepsilon, 0.4]) = 0.4^2 - (0.4 - \varepsilon)^2 \]
We will focus mostly on two kinds of random variables:

**discrete:** There are finitely (or countably) many possible values \( \{k_1, k_2, k_3, \ldots \} \) for \( X \).

\( \mu_X \) is described by the probability mass function \( p_X(k) = P(X = k) \) for each \( k \in \{k_1, k_2, k_3, \ldots \} \).

In this case, by the laws of probability,

\[
p_X(k) \geq 0 \quad \text{for each } k, \quad \sum_{j=1}^{\infty} p_X(j) = 1.
\]

**continuous:** For any real number \( t \in \mathbb{R} \), \( P(X = t) = 0 \).

\( \mu_X \) is captured by understanding \( P(X \leq r) \) as a function of \( r \).

Eg.

\[
P(X \in [a, b]) = P(\{X = a\} \cup \{X \in (a, b]\})
\]

\[
= P(X = a) + P(X \in (a, b])
\]

\[
= P(X \leq b) - P(X \leq a)
\]
**Cumulative Distribution Function (CDF)**

For any random variable $X$, $F_X(r) = P(X \leq r)$.

Eg. $M_X = \text{Bin}(3, \frac{1}{2})$

<table>
<thead>
<tr>
<th>$k$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_X(k)$</td>
<td>$\frac{1}{8}$</td>
<td>$\frac{3}{8}$</td>
<td>$\frac{3}{8}$</td>
<td>$\frac{1}{8}$</td>
</tr>
</tbody>
</table>

Actually works in all cases — including discrete.

- $r < 0$, $\{X \leq r\} = \emptyset \quad P(X \leq r) = 0$
- $0 \leq r < 1$, $\{X \leq r\} = \{X = 0\} \quad P(X \leq r) = P(X = 0) = \frac{1}{8}$
- $1 \leq r < 2$, $\{X \leq r\} = \{X = 0, 1\} \quad P(X \leq r) = P(X = 0) + P(X = 1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$
- $2 \leq r < 3$, $\{X \leq r\} = \{X = 0, 1, 2\} \quad P(X \leq r) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$
- $r \geq 3$, $\{X \leq r\} = \{X = 0, 1, 2, 3\} \quad P(X \leq r) = 1$
Properties of the CDF $F_X(r) = \mathbb{P}(X \leq r)$

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. 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 continuous function.
**Densities**

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

\[ X \text{ discrete, } \in \{k_1, k_2, k_3, \ldots \} \]

\[ p_X(k) = \Pr(X = k) \]

\[ \Pr(X \in A) = \sum_{k \in A} p_X(k) = \sum_{k \in A} p_X(k) \]

\[ p_X(k) \geq 0, \quad \sum_k p_X(k) = 1. \]

\[ X \text{ continuous} \]

\[ \Pr(X = t) = 0 \text{ for all } t \in \mathbb{R}. \]

\[ \Pr(X \in A) = \int f_X(t) \, dt \]
Eg. Shoot an arrow at a circular target of radius 1.

\[ Y = \text{distance from center.} \]

\[ \int_{-\infty}^{r} f(t) \, dt \]

\[ = \mathbb{P}(Y \in (-\infty, r]) = F_Y(r) = \begin{cases} 
0, & r < 0 \\
r^2, & 0 \leq r < 1 \\
1, & r \geq 1 
\end{cases} \]