Math 180A: Intro to Probability
(for Data Science)

www.math.ucsd.edu/~tkemp/180A

Today: § 3.2, 2.4
Next: § 2.5, 4.4

Lab 2.2 due TONIGHT!
HW3 due Friday.
Midterm next Wed evening

* Regrade requests for HW1:
  window Tuesday 10/15
  8am - 11pm

* Regrade requests for Lab1:
  Thursday 10/17

* separate request for each problem
  detailed, polite responses please.

* HW3 Problem 8: changed 2.25 → 2.21
**Cumulative Distribution Function (CDF)**

For any random variable $X$, $F_X(r) = P(X \leq r), \quad r \in \mathbb{R}$

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: $\lim_{t \to r^+} F_X(t) = F_X(r)$.

---

**Discrete random variable**: finite or countable set of values $t_1, t_2, t_3, \ldots$ with $P(X = t_j) > 0$ and $\sum_j P(X = t_j) = 1$.

**Continuous random variable**: for each real number $t$, $P(X = t) = 0$. Because of (1) & (3) above, this implies that $F_X$ is continuous.
Densities

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

$X$ discrete, $\epsilon \{t_1, t_2, t_3, \ldots \}$

$p_X(t) = P(X = t)$ probability mass function

$P(X \in A) = \sum_{t \in A} P(X = t)$

$= \sum_{t \in A} p_X(t)$

$p_X(t) \geq 0$, $\sum_{t} p_X(t) = 1$.

$X$ continuous

$P(X = t) = 0$ for all $t \in \mathbb{R}$.

But maybe there is an "infinitesimal" prob. mass function $f_X$.

$\Rightarrow P(X \in A) = \int_{A} f_X(t) \, dt$

i.e., $A = (-\infty, r]$

$P(X \leq r) = \int_{-\infty}^{r} f_X(t) \, dt$

$P(X \in [a, b]) = \int_{a}^{b} f_X(t) \, dt$

$\int_{-\infty}^{\infty} f_X(t) \, dt = 1$, $f_X(t) \geq 0$. 
Eg. Shoot an arrow at a circular target of radius 1.

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

\[
\int_{-\infty}^{r} f(t) \, dt \quad ?? \quad P(Y \in (\infty, r]) = F_Y(r) = \begin{cases} 0, & r \leq 0 \\ r^2, & 0 \leq r < 1 \\ 1, & r \geq 1 \end{cases}
\]

"Solve for f"

\[
\frac{1}{\pi} \int_{0}^{1} \frac{1}{r} \, dr = \int_{1}^{\infty} f(r) \, dr
\]

\[
\Rightarrow F(r) = \begin{cases} 0, & 0 < r \leq 1 \\ r^2, & r > 1 \end{cases}
\]

\[
f(r) = \begin{cases} 0, & 0 < r \leq 1 \\ 2r, & r > 1 \end{cases}
\]

\[
P(Y \in [0.1, 0.2]) \cup [0.9, 1]) = \int_{0.1}^{0.2} 2r \, dr + \int_{0.9}^{1} 2r \, dr
\]
Theorem: If $F_X$ is continuous and piecewise differentiable, then $X$ has a density $f_X = F_X'$. 

Proof: FTC. \[ \square \]

Eg. Let $X$ = a uniformly random number in $[0,1]$. As we discussed in lecture 2, this means

$$P(X \in [s,t]) = t - s \quad \text{if} \quad 0 \leq s < t \leq 1.$$

$$F_X(r) = P(X \leq r) = \begin{cases} 0 & r \leq 0 \\ r - 0 & 0 \leq r \leq 1 \\ 1 & r \geq 1 \end{cases}$$

$$f_X(r) = F_X'(r) = \begin{cases} 0 & r < 0 \\ 1 & 0 \leq r \leq 1 \\ 0 & r > 1 \end{cases}$$

$X \sim \text{Unif}(0,1)$

$Z \sim \text{Unif}(a,b) \Rightarrow f_Z(t) = \begin{cases} 0 & a \leq t < 0 \\ \frac{1}{b-a} & 0 \leq t \leq b \\ 0 & b > t \end{cases}$
Eg. Let \( f(t) = c \sqrt{b^2-t^2} \) for \( |t| \leq b \), 0 otherwise (for some positive constants \( b, c > 0 \)).

Is \( f \) a probability density?

- \( f \geq 0 \) \( \checkmark \)
- \[ \int_{-\infty}^{\infty} f(t) \, dt = \int_{-b}^{b} c \sqrt{b^2-t^2} \, dt = c \int_{-b}^{b} \sqrt{b^2-t^2} \, dt \]
  \[ \text{Subs: } t = bs. \quad c \int_{-1}^{1} \sqrt{b^2-(bs)^2} \, b \, ds = c b \int_{-1}^{1} \sqrt{b^2(1-s^2)} \, ds \]
  \[ = c b^2 \int_{-1}^{1} \sqrt{1-s^2} \, ds \]
  \[ = c b^2 \frac{\pi}{2} \]

\( \frac{c b^2 \pi}{2} = 1 \)

Must have \( \frac{c b^2}{2} = \frac{2}{\pi} \).

Eg. For any \( b \), \( c = \frac{2}{\pi b^2} \). \( \checkmark \)
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.

The random variable \( Z \) is

(a) continuous
(b) discrete
(c) neither
(d) both

\[
X \sim \text{Uniform}([100, 1500])
\]

\[
f_X(t) = \begin{cases} \frac{1}{1500-100} & 100 \leq t \leq 1500 \\ 0 & \text{otherwise} \end{cases}
\]

\[
P(Z=r) = \begin{cases} 1 & r < 500 \\ 0 & r \geq 500 \end{cases}
\]

\[
\begin{align*}
P(Z=500) &= P(Z \geq 500) \\
&= P(X \geq 500) \\
&= \int_{500}^{1500} \frac{1}{1500-100} \, dt \\
&= \frac{5}{7} > 0
\end{align*}
\]