Today: CDF and PDF

Next: ASV 2.4, 2.5, 4.4

Video: Prof. Todd Kemp, Fall 2019

Week 3:

- Homework 2 (due Friday, January 22)
- Midterm 1 (Wednesday, January 27) - Lectures 1-7
- Regrades for HW1: Mon, Jan 25 - Tue, Jan 26 (PST) on Gradescope
**Cumulative Distribution Function (CDF)**

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

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

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

3. The function \( F_X \) is right-continuous: \( \lim_{t \to r^+} F_X(t) = F_X(r) \).

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

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

\[ F(t_4) \quad \text{no jumps} \]
\[ F(t_5) \quad - - - - - - 0 \]
\[ \mathbb{P}(X \in (t_j, t_k]) = F(t_k) - F(t_j) \]
Some continuous random variables have probability densities. This is an infinitesimal version of a probability mass function.

- **Discrete**: \( X \) discrete, \( \epsilon \{ t_1, t_2, t_3, \ldots \} \)

  \[
  p_X(t) = P(X = t)
  \]

  \[
  P(X \in A) = \sum_{t \in A} P(X = t) = \sum_{t \in A} p_X(t)
  \]

  \[p_X(t) \geq 0, \quad \sum_{t} p_X(t) = 1.\]

- **Continuous**: \( X \) continuous

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

  **BUT**

  Maybe there is an "infinitesimal" prob. mass function \( f_X \).

  \[
  P(X \in A) = \int_{A} f_X(t) \, dt
  \]

  \[\text{i.e.,} \quad A = (-\infty, r] \]

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

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

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

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

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

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

"Solve for f" \[ \frac{d}{dr} \int_{-\infty}^{r} f(t) \, dt = \frac{d}{dr} \frac{1}{2} r^2 \quad \text{if } r \leq 1 \]

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

\[ f_Y(r) = \begin{cases} \frac{1}{2}, & 0 \leq r \leq 1 \\ 0, & r \geq 1 \end{cases} \]

\[ \mathbb{P}(Y \in [0.1, 0.2] \cup [0.9, 1]) = \int_{0.1}^{0.2} 2r \, dr + \int_{0.9}^{1} 2r \, dr \]

\[ 0.2 \left( \frac{0.2^2}{2} - \frac{0.1^2}{2} \right) + 1^2 - (0.9)^2 \]
Theorem: If $F_X$ is continuous and piecewise differentiable, then $X$ has a density $f_X = F_X'$. 

Proof: FTC. \[ \square \]

Eg. Let $X$ be a uniformly random number in $[0,1]$. As we discussed in lecture 2, this means $F_X(r) = \mathbb{P}(X \leq r) = \begin{cases} 0 & r \leq 0 \\ r - 0 & 0 < r \leq 1 \\ 1 & r > 1 \end{cases}$

$\therefore f_X(r) = \frac{d}{dr} F_X(r) = \begin{cases} 0 & r < 0 \\ 1 & 0 \leq r \leq 1 \\ 0 & r > 1 \end{cases}$

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

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

Is \( f \) a probability density?

\[ f \geq 0 \quad \checkmark \]

\[ \int_{-\infty}^{\infty} f(t) \, dt = \int_{-b}^{b} c\sqrt{b^2 - t^2} \, dt = c \int_{-b}^{b} \sqrt{b^2 - t^2} \, dt \]

Subs: \( t = bs \)

\[ = c b \int_{-1}^{1} \sqrt{b^2 - (bs)^2} \, bds \]

\[ = c b \int_{-1}^{1} \sqrt{1 - s^2} \, ds \]

\[ = c b^2 \int_{-1}^{1} \sqrt{1 - s^2} \, ds \]

\[ = c b^2 \frac{\pi}{2} \]

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

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

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

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

\[ \text{Pr}(Z=r) = 0, \quad r < 500 \]

\[ \text{Pr}(Z=500) = \text{Pr}(X \geq 500) = \int_{500}^{1500} \frac{1}{1400} \, dt = \frac{5}{7} > 0 \]