STAT 230 SECTION 2 LECTURE 22

Today's Agenda

Last time:

- More examples on variance
- Continuous random variables (motivation)

Today (Lec 22, 06/22):

- Continuous random variables:
  - definition
  - cdf and pdf
  - expected value and variance of continuous random variables
- the continuous uniform distribution

Chapter 8: Continuous Random Variables

Recall...

- a random variable is a function $X : S \to \mathbb{R}$.
- If the range $X(S)$ is...
  - countable or finite, X is called a discrete random variable
  - an interval $(a, b) \subseteq \mathbb{R}$, X is called a continuous random variable

Cutting a stick of length 2

- Suppose you cut a stick of length 2 at random and denote by X the cutting point. Then X can take values in $(0, 2)$.
- What is the probability that the cut X is
  - ... between 0 and 1?
  - ... between 1 and 2?
  - ... between 0 and 1/2?
  - ... smaller than 0?
  - ... larger than 1?
  - ... smaller than t for some $t \in (0, 2)$?
- The cdf $F(x) = P(X \leq x)$ is then

$$F(x) =$$

- And for any $x \in (0, 2)$,

$$P(X = x) =$$

So it seems that, for a continuous random variable, assigning probabilities to **intervals** is more “natural” than assigning probabilities to specific values.

We shall do exactly that by defining the cumulative distribution function for a continuous random variable.

Definition (CDF of continuous random variable)

The cumulative distribution function of a continuous random variable X, $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$, is a function such that

1. $F(x)$ is defined for all $x \in \mathbb{R}$,
2. $F(x)$ is a non-decreasing function of x for all $x \in \mathbb{R}$,
3. $\lim\limits_{x \to -\infty} F(x) = 0$ and $\lim\limits_{x \to \infty} F(x) = 1$,
4. $P(a < X \leq b) = F(b) - F(a)$,

Note: $P(X = x) = 0$ for all $x \in \mathbb{R}$!

Strict inequalities don't matter...

If X is a continuous random variable, then

$$P(a < X \leq b) = F(b) - F(a)$$

$$P(a \leq X \leq b) = P(a < X \leq b) + P(X = a) = F(b) - F(a) + 0$$

$$P(a < X < b) = P(a < X \leq b) - P(X = b) = F(b) - F(a) - 0$$

$$P(a \leq X < b) = P(a < X \leq b) + P(X = a) - P(X = b) = F(b) - F(a)$$

so if X is continuous, all these probabilities coincide!

If X was discrete, these 4 probabilities could all be different.

So we now have a tool for computing probabilities of a continuous random variable. However, the cdf doesn't show directly which values of X are more likely or less likely.

To find that out, let's see how the CDF changes as the value of x changes over a small interval $[x, x + \Delta x]$.

We have

$$P(x \leq X \leq x + \Delta x) = F(x + \Delta x) - F(x).$$

To study how CDF changes, we want to look at its rate of change as Δx becomes really small. That is,

$$\lim_{\Delta x \to 0} \frac{P(x \leq X \leq x + \Delta x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}.$$

Looks familiar? That's the derivative!

Definition (Probability density function)

The *probability density function (pdf)* of a continuous random variable X is the derivative of the cdf

$$f(x) = \frac{d}{dx} F(x),$$

where $f(x)$ is continuous.

Definition (Support of pdf)

The support of a pdf $f(x)$ is defined as

$$supp(f) = \{x \in \mathbb{R} : f(x) \neq 0\}.$$

Integrals of a pdf f over a domain D may be computed on $D \cap supp(f)$.

Properties of the pdf

a)

$$f(x) \geq 0 \quad \forall x \in \mathbb{R}$$

b)

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

c)

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx \quad \text{for } a < b \in \mathbb{R}$$

Example

Suppose that X is a continuous random variable with probability density function

$$f(x) = \begin{cases} cx(1-x) & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

a) Compute c so that this is a valid pdf
b) Graph $f(x)$
c) Compute $P(X \geq 1/2)$
d) Compute $P(1/4 \leq X \leq 3/4)$
e) Compute $P(X = 1/2)$

Example

Suppose that X is a continuous random variable with probability density function

$$f(x) = \begin{cases} x & \text{if } 0 \leq x \leq \sqrt{2}, \\ 0 & \text{otherwise} \end{cases}$$

a) Compute $F_X(x)$

b) Graph $F_X(x)$

c) Compute $P(1/2 \leq X < 1)$.

Definition (Expectation of continuous RV)

If X is a continuous random variable with pdf $f(x)$, and $g : \mathbb{R} \to \mathbb{R}$, then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

It follows that

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx.$$

$$Var(X) = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - E[X])^2 f(x) dx.$$

The shortcut formula for computing the variance still holds in the continuous case:

$$Var(X) = E[X^2] - (E[X])^2.$$

Example

Suppose X has pdf

$$f(x) = \begin{cases} 6x(1-x) & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise} \end{cases}$$

Compute $E[X]$ and $Var(X)$

Question

Suppose X has pdf $f(x)$, and f is an even function about the origin on $\mathbb{R}$ (i.e. $f(x) = f(-x)$). If $E[X]$ is well defined, which of the following is true:

A $E[X] < 0$

B $E[X] = 0$

C $E[X] > 0$

Example

Suppose X has CDF

$$F(x) = \begin{cases} 0 & x < 0, \\ \frac{x^2}{2} & 0 \leq x < 1/2 \\ \frac{7}{4}x - \frac{3}{4} & 1/2 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

Compute $E(X)$ and $Var(X)$.

So far, we have a way of dealing with the expectation of a function of random variable $g(X)$ (Law of Unconscious Statistician).

However, we may be interested in the distribution of $g(X)$ itself.

Example: $X =$ time until a light bulb breaks, $g(X) =$ the amount of work done before the light bulb breaks

If the function g has an inverse over the range of X, then we have a fairly easy way of obtaining the distribution of $Y = g(X)$.
In short, the method is as follows.

a) Write the CDF of Y as a function of X, i.e., start by $P(Y \leq y)$ and try to express it using X.
b) Use the CDF of X to find the CDF of Y. If you want the PDF of Y, then you can take the derivative of the CDF.
c) Find the range of Y.

Example

Let X be a continuous random variable with the following pdf and cdf:

$$f(x) = \begin{cases} \frac{1}{4} & 0 < x \leq 4, \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x}{4} & 0 < x < 4, \\ 1 & x \geq 4. \end{cases}$$

Find the pdf of $Y = X^{-1}$.

We now introduce the first continuous distribution.

Definition

We say that X has a continuous uniform distribution on (a, b) if X has pdf

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b), \\ 0 & \text{otherwise} \end{cases}$$

This is abbreviated $X \sim U(a, b)$.

Example

Let $X \sim U(a, b)$. Show the following.

a) $E(X) = \frac{a+b}{2}$

b) $Var(X) = \frac{(b-a)^2}{12}$

Example

Suppose that the angle measured from the principal axis to the point of a spinner is uniformly distributed on $[0, 2\pi]$. You win the prize you want if the point lands in $[\frac{3\pi}{4}, \frac{3\pi}{2}]$. Given that the point will stop in the top half of the circle, what is the probability that you win the prize you want.

Question

Suppose $X \sim U(0,1)$, and that $Y = \frac{2}{X} - 1$. What is the range of Y?

A $Y(S) = [0, \infty)$

B $Y(S) = [1, 3]$

C $Y(S) = [0, \infty)$

D $Y(S) = [0, 2]$

E $Y(S) = [1, \infty)$