Chapter 4 - Random Variables and Distributions

STAT 251

Lecture 9

Continuous Random Variables

Dr. Lasantha Premarathna

Chapter 4 - Learning Outcomes

- Notations
- Discrete Random Variables
- Continuous Random Variables
- Probability mass function (pmf)
- Probability density function (pdf)
- The Mean, the Variance and the Standard Deviation
- Cumulative Distribution Function (cdf)
- Max and Min of Independent Random Variables

Continuous Random Variable

The continuous random variables are used in relation with ‘continuous’ type of outcomes, as for examples

- the lifetime of a system or component
- pH values of chemical components
- in the study of the ecology of a lake, depth of a randomly chosen lake

Probability Density Function (pdf)

Mass : for discrete data.

- Let X be a continuous random variable. Then the probability density function (pdf) of X is a function $f(x)$ such that for any two numbers a and b with $a < b$

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

- $f(x)$ to be a legitimate pdf, it must satisfy
  1. $f(x) \geq 0$ for all x
  2. $\int_{-\infty}^{\infty} f(x)dx = 1$

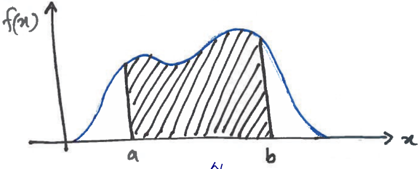
     $=$ this is the area under the graph of $f(x)$

IBP

$$\frac{df}{dx} \cdot g(x) + \frac{dg}{dx} \cdot f(x) = \frac{d}{dx}(g(x) \cdot f(x)) \qquad g(x) \cdot f(x) = f(x) \cdot \int g(x) + g(x) \cdot \int f(x)$$

Probability Density Function (pdf)

- Probability that X takes on a value in the interval [a,b] is the area above this interval and under the graph of the density function as given in the figure.



- Total area = 1
- Total prob = 1

Area under graph.

- Probability calculation

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

P at any pt = 0.

- Notice that, unlike in the discrete case, the inclusion or exclusion of the end points a and b doesn't affect the probability that the continuous variable X is in the interval.

Cumulative distribution function (cdf) of a continuous rv

pdf ⇒ $f(x)$

cdf ⇒ $F(x)$

X ↑ Prob.

x ↑ val

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)\,dx.$$

- The cumulative distribution function (cdf) of a continuous random variable X with pdf $f(x)$ is defined as

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\mathrm{d}t, \qquad \text{for all } x$$

Calculate probabilities - continuous rv

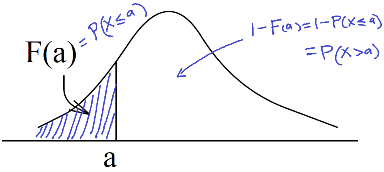
Note that in particular

- $P(X > a) = 1 - F(a)$

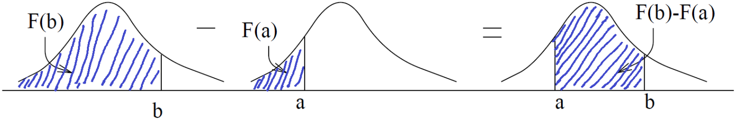
-reminder, total $p = 1$

$P(X > a) = 1 - P(X \leq a)$

$= 1 - F(a)$



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



probability on (a, b) under density function $f(x)$

pdf & cdf

$f(x) = F'(x) = \frac{dF(x)}{dx}$

pdf.

$F(x) = \int_{-\infty}^{a} f(x)\, dx$

- By fundamental theorem of calculus

$$f(x) = F'(x) = \frac{dF(x)}{dx}, \qquad \text{for all } x$$

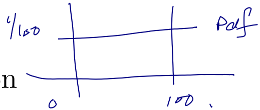
- We can go back and forth from the probability density function to the cumulative distribution function and vice versa using

$f(x) \longrightarrow F(x)$ by integration the pdf pdf $\rightarrow$ cdf <Integr.>

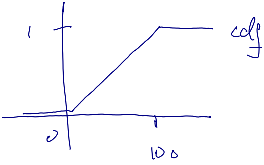
$F(x) \longrightarrow f(x)$ by differentiation the cdf cdf $\rightarrow$ pdf. <diff.>

Example 6

Consider a random variable X with density function



$$f(x) = \begin{cases} \dfrac{1}{k} & ; 0 \leq x \leq 100 \\ 0 & ; \text{otherwise} \end{cases}$$



(a) Calculate the value of k and then the cdf

(b) Calculate $P(40 \leq X \leq 70)$

$\int_0^{100} \frac{1}{k} = 1$

$= \frac{1}{k} \times x \Big|_{x=0}^{100} = \frac{1}{k} \times 100 - \frac{1}{k}(0) = 1$

$\frac{100}{k} = 1$

$k = 100$ //

$$\text{cdf} = \begin{cases} 0 & : x < 0 \\ \frac{x}{100} & : 0 < x \leq 100 \\ 0 & : x > 100 \end{cases}$$

cdf always incr.

(b) $\int_{40}^{70} \frac{1}{100}\, dx$.

$= \frac{1}{100} \times x \Big|_{40}^{70}$

$= \frac{70}{100} - \frac{40}{100} = \frac{30}{100} = 3/10$ //

Find Median & IQR of the rv X

- Find $F(x)$
- To find the median, solve x such that $F(x) = 0.5$
- To find the Q_1, solve x such that $F(x) = 0.25$
- To find the Q_3, solve x such that $F(x) = 0.75$

$$\Rightarrow IQR = Q_3 - Q_1$$

Mean of a Continuous random variable

- Mean/Expectation /Expected value of a continuous random variable X with pdf $f(x)$ is

Pop. mean

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

expected val.

- in general

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

$g(X)$ is a function of X

Variance and Standard deviation of a continuous rv

Consider the continuous random variable X with pdf $f(x)$

- Variance of X is

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

- Standard deviation of X is

$$\sigma = \mathrm{SD}(X) = \sqrt{\mathrm{Var(X)}}$$

** Following formula is often used to calculate the variance

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2$$

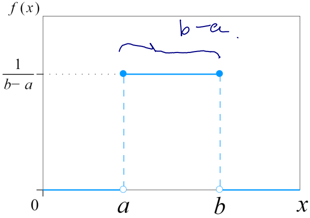
Some continuous Probability distributions

Uniform Distribution –

If X is a uniform random variable, X is uniformly distributed on the interval [a,b]

pdf of X is

$$f(x) = \begin{cases} \dfrac{1}{b-a} & ; a \le x \le b \\ 0 & ; \text{otherwise} \end{cases}$$



Notation: $X \overset{\text{follows}}{\sim} U(a,b) \leftarrow$ uniform distr.

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

$a = 1 = b-a \times \frac{1}{b-a}$ b − a

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

(obtain μ and σ^2)

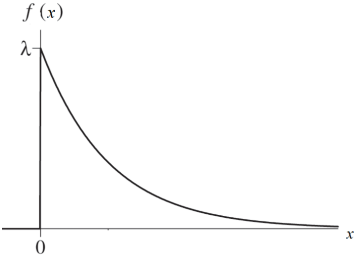
$\int_{-\infty}^{\infty} x f(x)\, dx$

Some continuous Probability distributions

Exponential Distribution

Exponential random variables are often used to model the time until an even occur. If X is a exponential random variable with $\lambda > 0$ (rate parameter), then the pdf of X is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ; x \geq 0 \\ 0 & ; x < 0 \end{cases}$$



Notation: $X \sim Exponental(\lambda)$

or $X \sim Exp(\lambda)$

$$\mu = E(X) = \frac{1}{\lambda}$$

$$\sigma^2 = Var(X) = \frac{1}{\lambda^2}$$

(obtain μ and σ^2)

Example 7

cdf $= \int_{-\infty}^{x} \frac{3}{8} x^2 dx$

$= \frac{3}{8} \frac{x^3}{3} = \frac{1}{8} x^3$

$= \begin{cases} 0 & x<0 \\ x^3/8 & 0<x<2 \\ 0 & x<2 \end{cases}$

(b) $P(1<x<2)$

$\int_1^2 \frac{3}{8} x^2 dx$

$= \frac{2^3}{8} - \frac{1^3}{8} = 1 - \frac{1}{8} = 7/8.$

Consider the pdf of the random variable X

$$f(x) = \frac{3}{8}x^2 \qquad ; 0 \leq x \leq 2$$

(a) Find the cdf of X

(b) Calculate $P(1 < X < 2)$

(c) Find the median of the distribution

(d) Find the mean and variance

(c) Median

$= \int_0^2 x \cdot \frac{3}{8} x^2$

$= \int_0^2 x^3 \cdot \frac{3}{8}$

$= \frac{3 x^4}{32} \Big|_0^2$

$= \frac{32}{32} - 0 = 1$

Before the next class ...

Visit the course website at canvas.ubc.ca

- Review the lecture 9 and related sections in the text book
- Topic of next class: **Properties of Mean and Varience and Covariance, Sum of Independent Random Variables, and Maximum and Minimum of Independent Random Variables**