Sri Lanka Institute of Information Technology

Faculty of Computing

IT2120 - Probability and Statistics

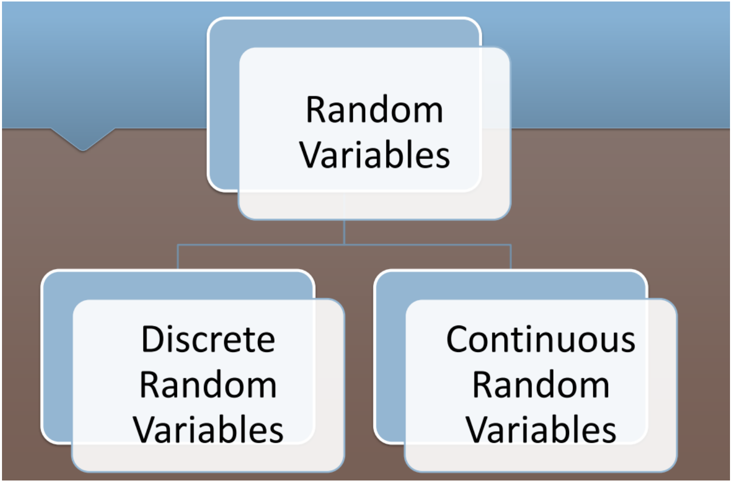
Ms. K. G. M. Lakmali

Year 02 and Semester 01

Lecture 6

CONTINUOUS PROBABILITY DISTRIBUTIONS

Random Variables



Continuous Random Variables

- A random variable is said to be continuous, if it can take any value within a range.
- Continuous data are frequently measured in some way rather than counted.
- If X is a continuous random variable, $Pr(X=a) = 0$ for any value of a.

Examples

- Temperature
- Heart beat of a patient
- Rainfall
- Waiting time for a bus

PROBABILITY DISTRIBUTIONS

- For continuous random variables, the probability distribution cannot be presented in a tabular form.
- Probability distribution function of a continuous random variable is known as probability density function (pdf).
- The area under the p.d.f. gives probability values

PDF - DEFINITION

The function $f_X(x)$ is a probability density function for the continuous random variable X, defined over the set of real numbers ($\mathbb{R}$), if

- $f_X(x) \geq 0$, for all $x \in \mathbb{R}$
- $\int_{-\infty}^{\infty} f_X(x)dx = 1$
- $\Pr(a < X < b) = \int_a^b f_X(x)dx$

Properties

Let X be a continuous random variable with a p.d.f. ($f_X(x)$), defined over the set of real numbers ($\mathbb{R}$).

- The c.d.f. $F_X(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f_X(x)dx$
- $E[g(x)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx$
- $V[g(x)] = E[g(x)^2] - \{E[g(x)]\}^2$

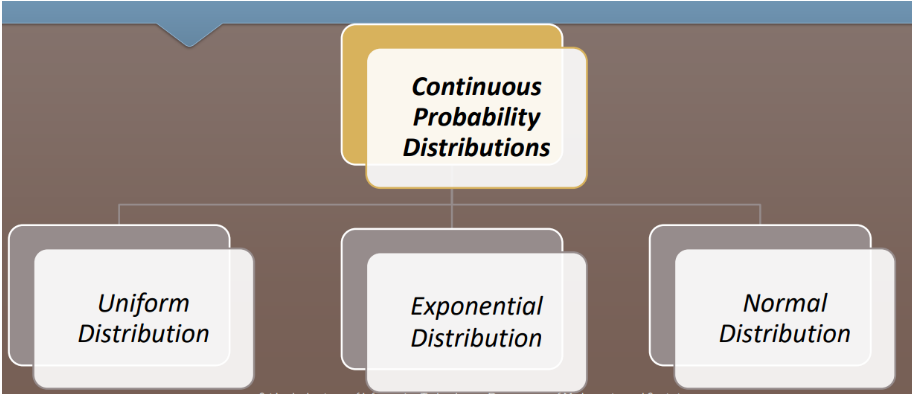
PDF - Example

Suppose that the error in the reaction temperature, in °C, for a controlled laboratory experiment is a continuous random variable X having the probability density function,

$$f_X(x) = \begin{cases} cx^2 & -1 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

1. Find the value of c
2. Find $\Pr(0 < X \leq 1)$.
3. Find the expected value and the variance.
4. Find the c.d.f.

Continuous Probability Distributions



Uniform Distribution

- This is one of the simplest probability distributions in statistics.
- Let X is a continuous random variable taking values in the range of a and b follows a uniform distribution. Then it can be represented as follows:

$$X \sim U(a, b)$$

- With the probability density function given by:

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

Uniform Distribution

Then, expected value and variance for X will be calculated as follows:

- $E(X) = \frac{b+a}{2}$
- $V(X) = \frac{(b-a)^2}{12}$

A uniform distribution where $a = 0$ and $b = 1$, called as **standard uniform distribution**. Here, $E(X) = \frac{1}{2}$ and $V(X) = \frac{1}{12}$.

Uniform Distribution - Example

1) It is given that arrival time of an elevator follows uniform distribution where it takes 0 to 50 seconds to come from one floor to another floor.

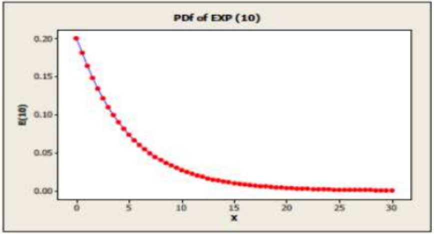
1. What is the probability that this elevator arrives within 30 seconds?
2. Find the expected arrival time of the elevator?
3. What is the variance?

Continuous Distributions

Exponential Distribution	Normal Distribution / Gaussian Distribution
The distribution is usually used to ***model life times***. (There is a link to the Poisson distribution)	This is most commonly used distribution. This is bell shaped distribution and perfectly symmetric around μ.
$X \sim Exp(\lambda)$	$X \sim N(\mu, \sigma^2)$
$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < \infty$
$E(X) = 1/\lambda$	$E(X) = \mu$
$V(X) = 1/\lambda^2$	$V(X) = \sigma^2$

Exponential Distribution

- Widely used in waiting line (or queuing) theory to model the length of time between arrivals in process.
- **Examples**: duration between two customers at Bank ATMs, To model patients entering to an accident ward



Exponential Distribution - Example

1) The time, in hours, during which an electrical generator is operational is a random variable that follows an exponential distribution with a mean of 160. What is the probability that a generator of this type will be operational for,

1. Less than 40 hours?
2. Between 60 and 160 hours?
3. More than 200 hours?

Standard Normal Distribution

- Normal distribution with $\mu = 0$ and $\sigma^2 = 1$ is known as the **Standard Normal Distribution**.
- Evaluating probabilities with Normal requires complex integration.
- To simplify the procedure, statistical tables are defined.
- But, tables for each combination of μ and σ^2 cannot be created.
- So, tables are only for the standard normal distribution.

Normal $\rightarrow$ Standard Normal

If $X \sim N(\mu, \sigma^2)$, Then

$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

Normal Distribution - Examples

1) For Z $\sim$ N(0, 1), calculate Pr(Z $\geq$ 1.13).

2) For X $\sim$ N(5, 4), calculate Pr($-2.5 < X < 1.13$).

3) The actual marks for FCS of Metro students revealed that they were normally distributed with a mean mark of 45 and a standard deviation of 22. What is the probability that a randomly chosen student will pass? (Assume that pass mark is 45)

Approximating Binomial Probabilities

Normal Distribution

- For $X \sim \text{Bin}(n, p)$ this approximation can be used if n is large and p is moderate.
- A general rule can be defined as, np and n(1-p) is greater than 5.
- Can be approximated with a r.v. with a distribution $\text{N}(\text{np}, \text{np}(1 - \text{p}))$.
- A continuity correction is needed because a discrete distribution is approximated with a continuous distribution.

Continuity Correction

If $X \sim \text{Bin}(n, p)$ is approximated with a r.v. $Y \sim N(np, np(1-p))$,

- $\Pr(X \leq a) = P(Y < a+0.5)$
- $\Pr(X \geq a) = P(Y > a-0.5)$
- $\Pr(X < a) = P(Y < a-0.5)$
- $\Pr(X > a) = P(Y > a+0.5)$
- $\Pr(X = a) = P(a-0.5 < Y < a+0.5)$

Example

Suppose that a sample of n = 1,600 tires of the same type are obtained at random from an ongoing production process in which 8% of all such tires produced are defective. What is the probability that in such a sample 150 or fewer tires will be defective?

Approximating Poisson Probabilities

Normal Distribution

- If $X \sim \text{Poisson}(\lambda)$ then if λ is greater than 20, the approximation can be used.
- Can be approximated with a r.v. with a distribution $\text{N}(\lambda, \lambda)$.
- A continuity correction is needed because a discrete distribution is approximated with a continuous distribution (just as in the case of the Binomial to Normal approximation).

Example

The annual number of earthquakes registering at least 2.5 on the Richter Scale and having an epicenter within 40 miles of down town Memphis follows a Poisson distribution with mean 22.5. What is the probability that at least 25 such earthquakes will strike next year?


THANKS!
Any questions?