Chapter 4: Continuous Random Variables and Probability Distributions

LEARNING OBJECTIVES

1. Continuous random variable:

(a) Probability density function

(b) Cumulative distribution function

(c) Mean and Variance

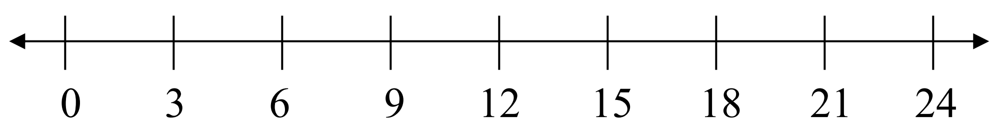
2. Common distribution: uniform and normal

3. Normal approximation to the Binomial and Poisson.

Definition

A continuous random variable is a random variable whose possible values includes in an interval of real numbers.

Hours spent studying in a day



The time spent studying can be any number between 0 and 24.

Definition

The **probability density function (pdf)** of a continuous random variable X is a function such that

(1) $f(x) \geq 0 \quad \forall \; x$

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

(3) $P(a \leq X \leq b) = \int_a^b f(x)dx$ for any a and b.

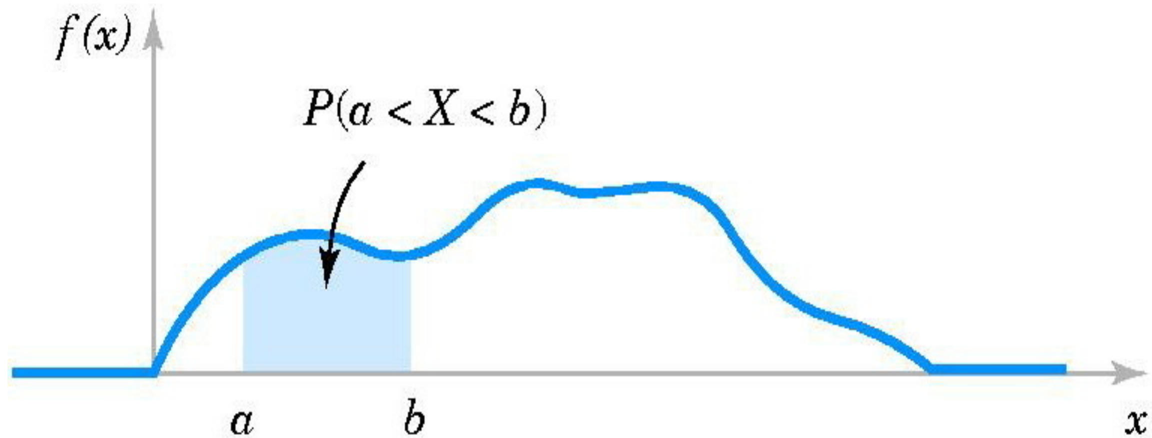


Figure 4-2 Probability determined from the area under $f(x)$.

Property

If X is a continuous random variable then for any x_1 and x_2 we have

$$\begin{aligned} P(x_1 \le X \le x_2) &= P(x_1 \le X < x_2) \\ &= P(x_1 < X \le x_2) \\ &= P(x_1 < X < x_2) \end{aligned}$$

Example

Let the continuous random variable X denote the diameter of a hole drilled in a sheet metal component. The target diameter is 12.5 millimeters. Most random disturbances to the process result in larger diameters. Historical data show that the distribution of X can be modeled by a probability density function

$$f(x) = 20e^{-20(x-12,5)}, \; x \geq 12.5$$

(a) If a part with a diameter larger than 12.60 millimeters is scrapped, what proportion of parts is scrapped?

(b) What proportion of parts is between 12.5 and 12.6 millimeters?

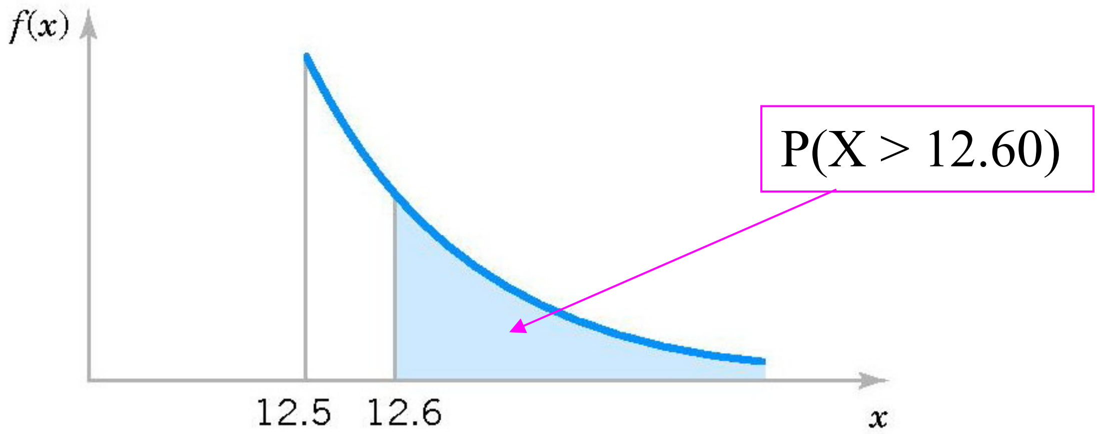
Solution

(a) $$P(X > 12.60) = \int_{12.6}^{+\infty} f(x)dx = 20\int_{12.6}^{+\infty} e^{-20(x-12.5)}dx$$

$$= -e^{-20(x-12.5)} \Big|_{12.6}^{+\infty} = 0.135$$

(b) $$P(12.5 < X < 12.6) = \int_{12.5}^{12.6} 20e^{-20(x-12.5)}dx$$

$$= -e^{-20(x-12.5)} \Big|_{12.5}^{12.6} = 0.865$$



Definition

The **cumulative distribution function (cdf)** of a continuous random variable X is

$$F(x) := \int_{-\infty}^{x} f(t)dt$$

for -∞ < x < +∞.

Let us return to the above example

$$f(x) = \begin{cases} 20e^{-20(x-12,5)} & x \geq 12.5 \\ 0 & x < 12.5 \end{cases}$$

$$\Longrightarrow \quad F(x) = \begin{cases} 1 - e^{-20(x-12,5)} & x \geq 12.5 \\ 0 & x < 12.5 \end{cases}$$

Definition

Suppose X is a continuous random variable with probability density function $f(x)$.

The **mean** or **expected value** of X is defined by

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

The **variance** of X is defined by

$$\sigma^2 = V(X) := \int_{-\infty}^{+\infty} (x-\mu)^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 f(x) dx - \mu^2$$

The **standard deviation** of X is $\sigma = \sqrt{\sigma^2}$.

Example

Assume that X is a continuous random variable with the following probability density function

$$f(x)=\begin{cases}20e^{-20(x-12,5)} & x\geq 12.5\\ 0 & x<12.5\end{cases}$$

Mean: $$EX=\int_{12.5}^{+\infty} xf(x)dx=\int_{12.5}^{+\infty} x20e^{-20(x-12.5)}dx$$

Integration by parts can be used to show that

$$EX=\left.\left(-xe^{-20(x-12.5)}-\frac{e^{-20(x-12.5)}}{20}\right)\right|_{12.5}^{+\infty}=12.55$$

Variance:

$$V(X) = \int_{12.5}^{+\infty} x^2 f(x)dx - (EX)^2 = 0.0025$$

Expected Value of a Function of a Continuous Random Variable

$$\boxed{E\,h(X) = \int_{-\infty}^{+\infty} h(x) f(x) dx}$$

Continuous uniform random variable over interval [*a*, *b*]

pdf:

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

Mean and Variance:

$$\mu = EX = \frac{a+b}{2} \quad , \quad \sigma^2 = V(X) = \frac{(b-a)^2}{12}$$

cdf:

$$F(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x < b \\ 1 & b \ge b \end{cases}$$

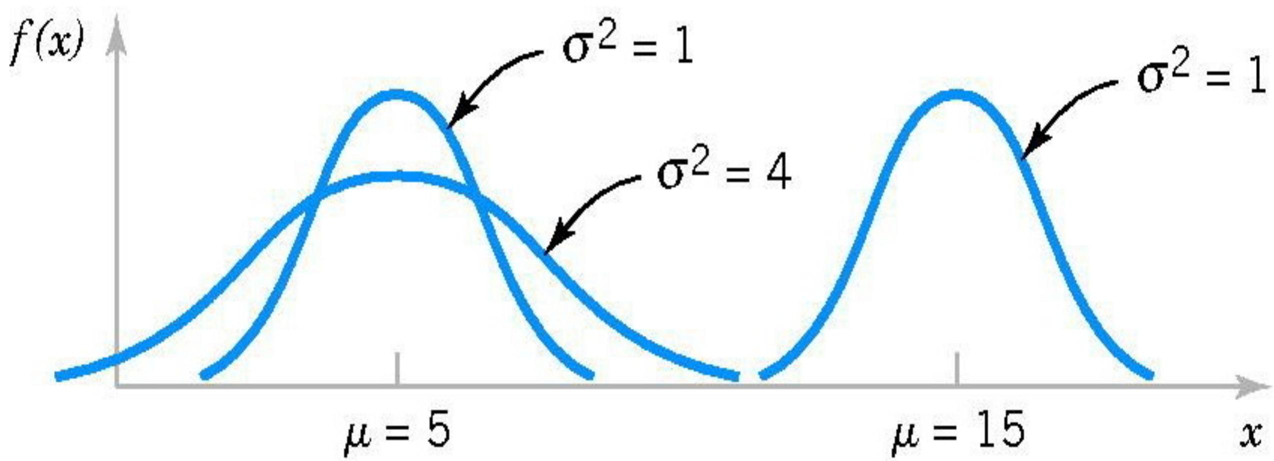
Normal random variable X ~ $N(\mu, \sigma^2)$

pdf:

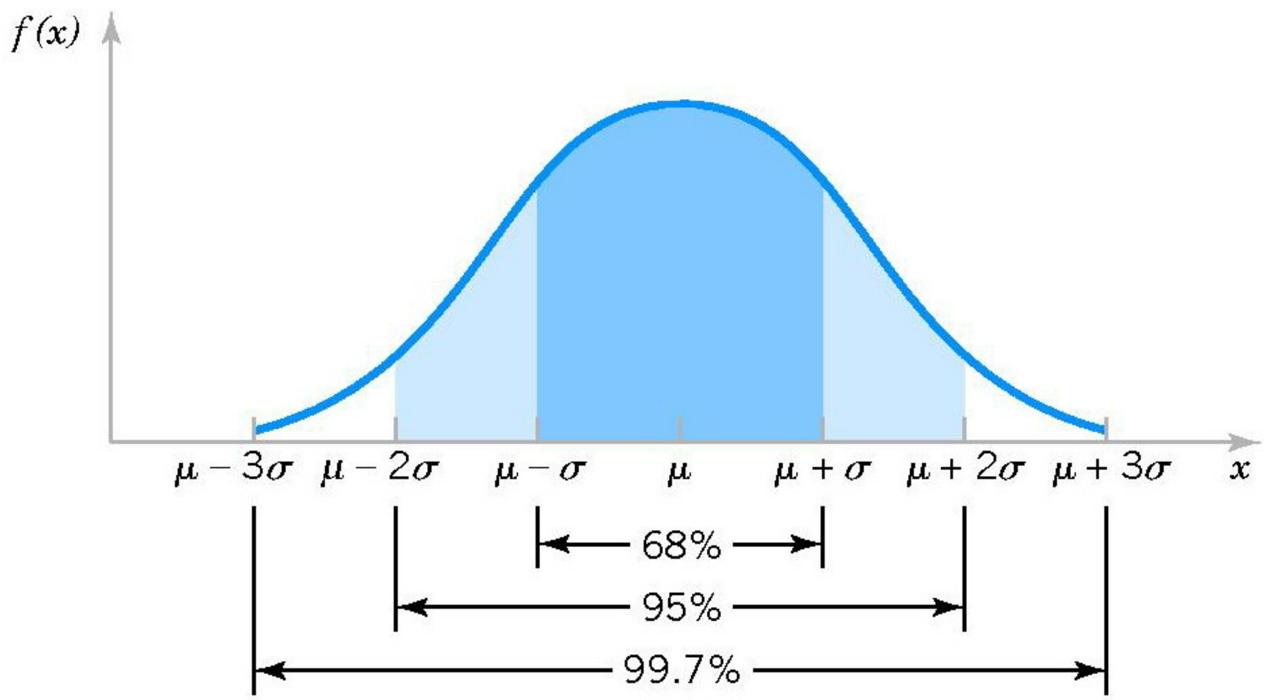
$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\ ,\ -\infty<x<+\infty$$

Mean and Variance:

$$E(X) = \mu \qquad V(X) = \sigma^2$$



3σ-rule


$f(x)$
$\mu - 3\sigma$
$\mu - 2\sigma$
$\mu - \sigma$
μ
$\mu + \sigma$
$\mu + 2\sigma$
$\mu + 3\sigma$
x
68%
95%
99.7%

Standard Normal Random Variable $Z \sim N(0,1)$

pdf:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \quad , \quad -\infty < x < +\infty$$

cdf:

$$\Phi(z) = \int_{-\infty}^{z} \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx.$$

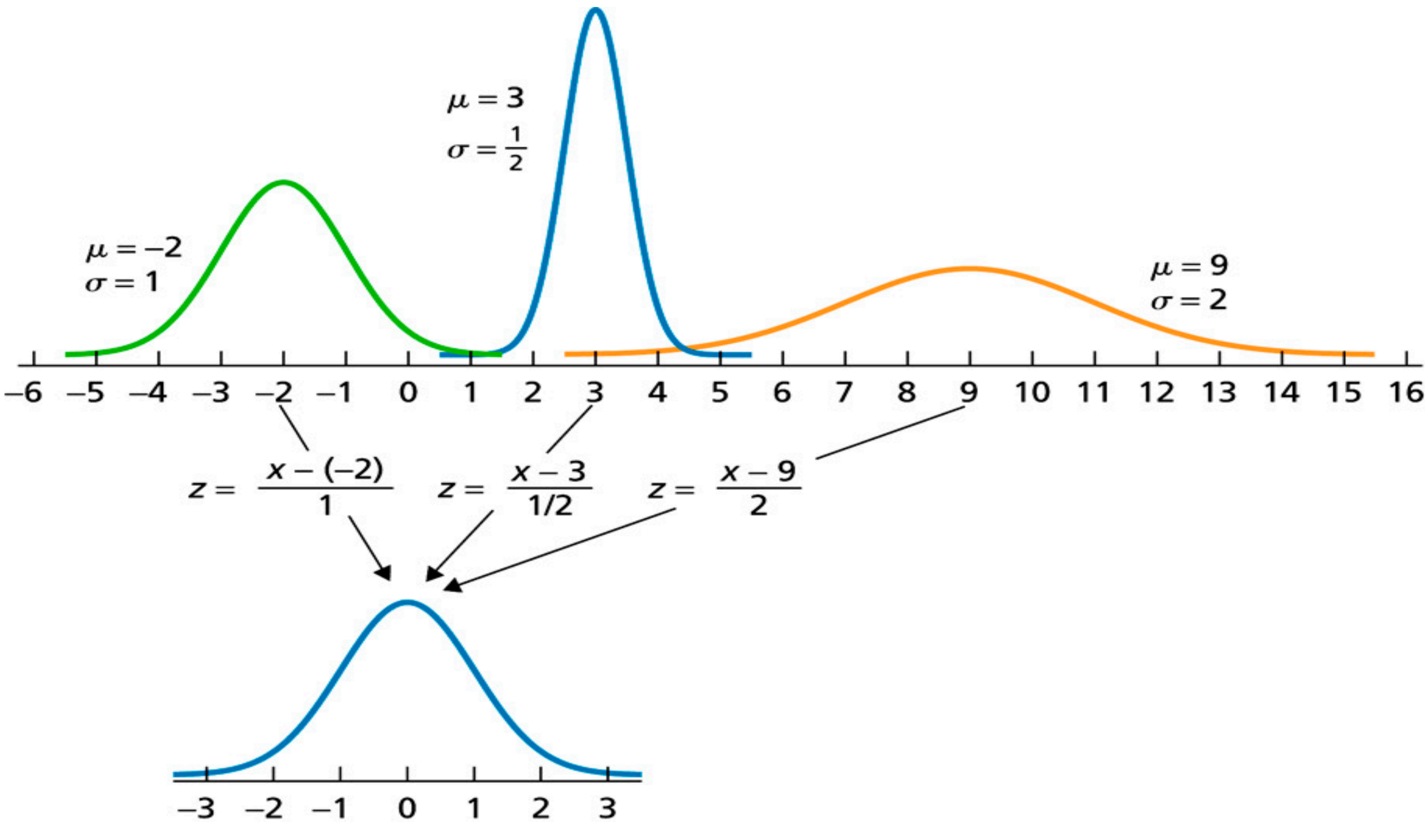
Standardizing

If X is a normal random variable $X \sim N(\mu, \sigma^2)$ then

$$Z = \frac{X-\mu}{\sigma}$$

is a standard normal random variable $N(0, 1)$.

Standardizing



Finding Probabilities

Problem 1: x is given known, find $P(X < x)$.

Problem 2: $P(X < x) = p$ is given known, find x.

Using Excel

1. $Z \sim N(0,1)$

To find $P(Z < z)$ when given z: =NORMSDIST(z)

To find z when $P(Z < z) = p$: =NORMSINV(p)

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

To find $P(X < x)$ when given x: =NORMDIST(x, μ, σ,1)

To find x when $P(X < x) = p$: =NORMINV(p, μ, σ).

Example

(a) Let $X \sim N(34, 144)$. Find $P(X < 43)$ and $P(24 < X < 37)$.

(b) Let $Z \sim N(0,1)$. Find the value of z to $P(Z > z) = 0.95$

(a) $P(X < 43)$ = NORMDIST(43, 34, 12,1) = 0.7734

$P(24 < X < 37) = P(X < 37) – P(X < 24)$

$= 0.5987 – 0.2023 = 0.3964$

(b) $P(Z < z) = 1 – P(Z > z) = 1 – 0.95 = 0.05$

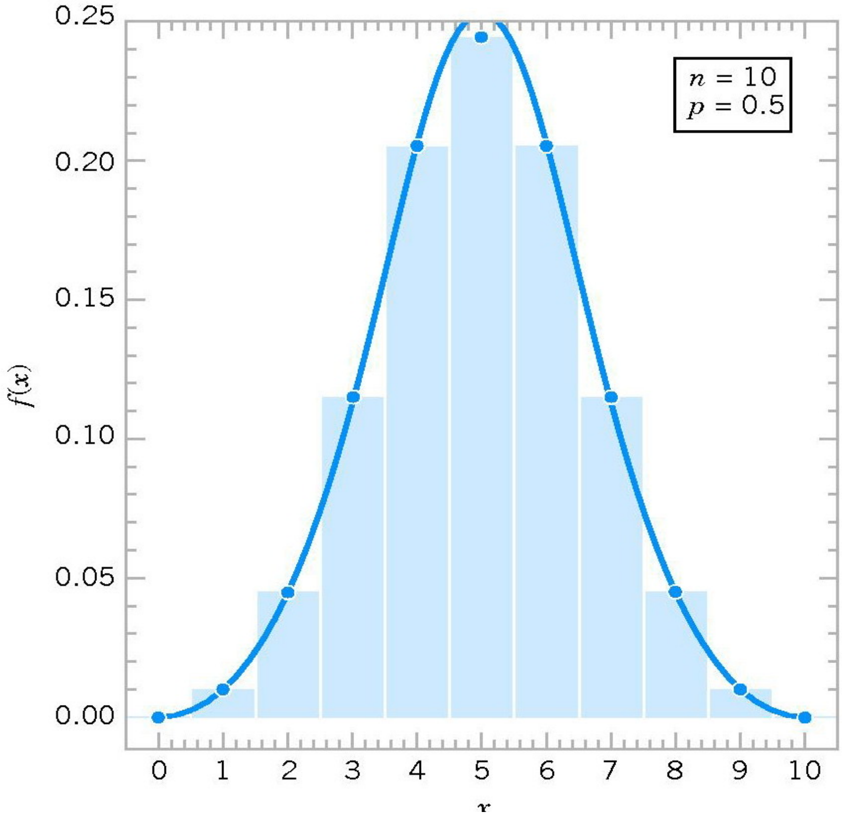
z =NORMSINV(0.05) = -1.65

Finding Probabilities: Using Table II

Table II Cumulative Standard Normal Distribution (*continued*)

z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373
0.8	0.788145	0.791030	0.793892	0.796731	0.799546	0.802338
0.9	0.815940	0.818589	0.821214	0.823815	0.826391	0.828944
1.0	0.841345	0.843752	0.846136	0.848495	0.850830	0.853141
1.1	0.864334	0.866500	0.868643	0.870762	0.872857	0.874928
1.2	0.884930	0.886860	0.888767	0.890651	0.892512	0.894350

Under certain conditions, the normal distribution can be used to approximate the binomial distribution and the Poisson distribution.



Example

Let $X \sim B(16\times10^6, 10^{-5})$. Find the probability $P(X > 150)$.

Solution

$$P(X > 150) = 1 - P(X \le 150)$$

$$= 1 - \sum_{x=0}^{150} C_{16\times10^6}^{x} (10^{-5})^x (1-10^{-5})^{16\times10^6 - x}$$

Clearly, this probability is difficult to compute.

Normal Approximation to the Binomial Distribution

If $X \sim B(n, p)$ then random variable

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

is approximately a standard random variable $N(0,1)$.

$$P(X \le x) \approx P\left(Z \le \frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right)$$

$$P(X \ge x) \approx P\left(Z \ge \frac{x - 0.5 - np}{\sqrt{np(1-p)}}\right)$$

Remark: The approximation is good for $np > 5$ and $n(1-p) > 5$.

Let us return to the above example

$$P(X > 150) = P(X \geq 151) = P(Z > \frac{150.5 - 160}{\sqrt{160(1 - 10^{-5})}})$$

$$= P(Z > -0.75) = 1 - P(Z < -0.75) = 0.773$$

Here, we use the result $np = 16\times 10^{6} \times 10^{-5} = 160$.

Normal Approximation to the Poisson Distribution

If X is a Poisson random variable $P(\lambda)$ then

$$Z = \frac{X-\lambda}{\sqrt{\lambda}}$$

is approximately a standard random variable $N(0,1)$.

$$P(X \le x) = P(Z \le \frac{x-\lambda}{\sqrt{\lambda}}) = \Phi(\frac{x-\lambda}{\sqrt{\lambda}}).$$

The approximation is good for **λ > 5.**

Example

Assume that the number of asbestos particles in a squared meter of dust on a surface follows a Poisson distribution with a mean of 1000. If a squared meter of dust is analyzed, what is the probability that less than 950 particles are found?

Let X = the number of asbestos particles in a squared meter of dust on a surface, then $X \sim P(1000)$

$$P(X \le 950) \cong P(Z \le \frac{950-1000}{\sqrt{1000}}) = P(Z \le -1.58) = 0.057$$

Continuous R.V
Pdf
Cdf
Mean & Variance
Common Distri.
Uniform
Normal
Normal approxi.
Summary

We have studied:

1. Continuous random variable:

(a) Probability density function

(b) Cumulative distribution function

(c) Mean and Variance

2. Common distribution: uniform and normal

3. Normal approximation to the Binomial and Poisson.

Homework: Read slides of the next lecture.