Lecture 3

May 13, 2019 4:49 PM

4.1 & 4.2 Continuous Random Variable and Probability Density Functions

Continuous. R. V.

A R.V. that can assume any value in a given interval.

Example: temperature, time.

uncountable infinite.

P.D.F

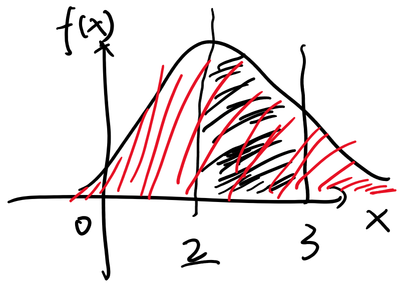
For a continuous R.V. X. A probability density function (p.d.f) is an function satisfies

(1) $f(X) \geq 0$

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

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

$\frac{1}{49C6}$ $\qquad \frac{1}{\infty} = 0$



$f(x=2) = P(X=2) = 0$

$P(2 \leq X \leq 3)$ = value/area.

$P(-\infty \leq X \leq +\infty) = 1$

Example:

Let X be a lifetime of an electrical component.

Suppose that $f(x) = ce^{-3x}$, $x > 0$

(a) Find C.

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

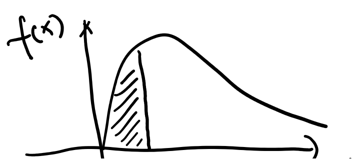
$\int_{-\infty}^{+\infty} ce^{-3x}\,dx = 1$

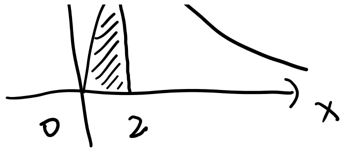
$\Rightarrow \int_0^{+\infty} ce^{-3x}\,dx = 1$

$\Rightarrow -\frac{1}{3}e^{-3x} \cdot c \Big|_0^{\infty} = 1$

$\Rightarrow c = 3$

(b) what proportion of component last less than 2 hours.





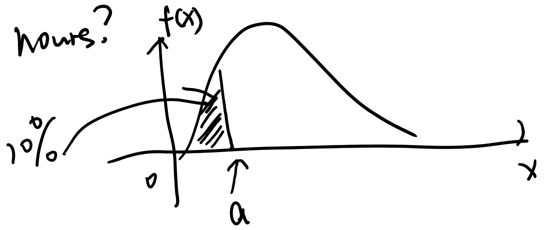
$$p(0 \le X \le 2) = \int_0^2 f(x)\,dx \qquad \text{(3rd property)}$$

$$= \int_0^2 3e^{-3x}\,dx$$

$$= -e^{-3x}\Big|_0^2$$

$$= 1 - e^{-6}$$

(C) 10% of such component last less than how many hours?



$$\int_0^a f(x)\,dx = 0.1$$

$$\int_0^a 3e^{-3x}\,dx = 0.1$$

$$1 - e^{-3a} = \frac{1}{10}$$

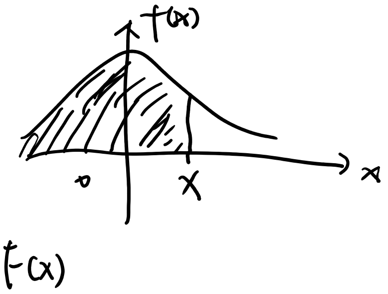
$$a = -\frac{1}{3}\ln\left(\frac{9}{10}\right)$$

4.3 Cumulative Distribution Function:

C.D.F.

The C.D.F of a continuous R.V. X is defined by

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



Example:

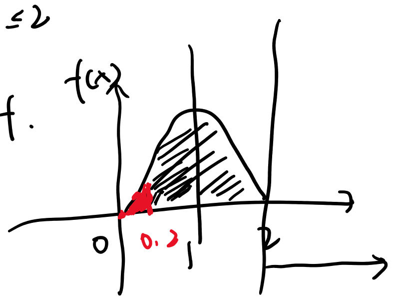
$$f(x) = \begin{cases} \frac{1}{4}(x+1), & 0 \le x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find C.D.F : $F(x)$ $\quad 0 \le X \le 2$

$$F(x) = \int_0^x f(x)\,dx \qquad 0 \le X \le 2$$

$$F(x) = \int_0^x f(x)\,dx \qquad 0 \le x \le 2$$
$$= \int_0^x \frac{1}{4}(x+1)\,dx$$
$$= \frac{1}{4}\left(\frac{1}{2}x^2 + x\right)\Big|_0^x$$
$$= \frac{1}{8}x^2 + \frac{x}{4}$$

p.d.f.



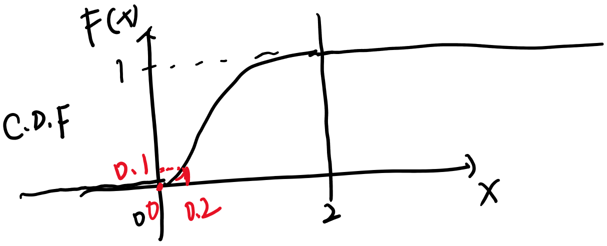
when $x > 2$

$F(x) = 1$

when $x < 0$

$F(x) = 0$

C.D.F

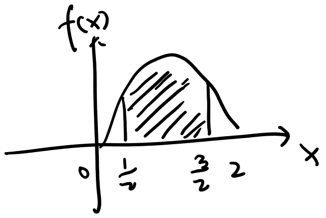


$$F(1) = P(0 \le X \le 1)$$
$$F(2) = P(0 \le X \le 2)$$

$P(0 \le X \le 0.2) = F(0.2) = 0.1$

(b) $P\left(\frac{1}{2} \le X \le \frac{3}{2}\right) = \int_{\frac{1}{2}}^{\frac{3}{2}} f(x)\,dx$ ① ✓

$= F\left(\frac{3}{2}\right) - F\left(\frac{1}{2}\right)$ ② ✓



Note: ① $P(a \le X \le b) = P(a < X < b) = F(b) - F(a)$

② $F(x) = \int_{-\infty}^{x} f(x)\,dx$

$\Leftrightarrow \dfrac{dF(x)}{dx} = f(x)$

4.4. Mean and Variance for Continuous R.V.

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

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

$$\sigma^2 = E(X^2) - [E(X)]^2$$

where $E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$

Example:

Suppose that the lifetime X has a p.d.f.

$$f(x) = \frac{1}{500} e^{-x/500}, \qquad x > 0$$

What's the average lifetime of X.

$$\mu = E(X) = \int_0^{+\infty} x f(x)\,dx$$

$\boxed{\int u\,dv = uv - \int v\,du}$

$$\mu = E(X) = \int_0^{+\infty} x f(x)\, dx$$

$$= \int_0^{\infty} \frac{1}{500} x e^{-x/500}\, dx$$

$$\boxed{\int u\,dv = uv - \int v\,du}$$

(Intergration by parts.

Let $u = x$, $dv = e^{-x/500}$

$du = dx$, $v = -500 e^{-x/500}$)

$$= \left[-500 x e^{-x/500}\right]\Big|_0^{\infty} + \int_0^{\infty} 500 e^{-x/500}\, dx \cdot \frac{1}{500}$$

$$= 500$$

$$E(X^2) = \int_0^{\infty} x^2 f(x)\, dx$$

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

4.6 The Normal Distribution.

Gaussian Distribution

Definition: X is a normal R.V. with parameter μ and σ if

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

Denote: $\boxed{X \sim N(\mu, \sigma^2)}$ $\qquad X \sim N(2,3) \Rightarrow \begin{cases} \mu = 2 \\ \sigma^2 = 3 \end{cases}$

$\therefore E(X) = \mu, \quad Var(X) = \sigma^2$

$1 + 2 + 3 \cdots + 100$

$$\begin{array}{ccccc} 1 & 2 & \cdots & 49 & 50 \\ 100 & 99 & & 52 & 51 \\ =101 & 101 & & 101 & 101 \end{array}$$

$\underbrace{\qquad\qquad\qquad}_{50}$

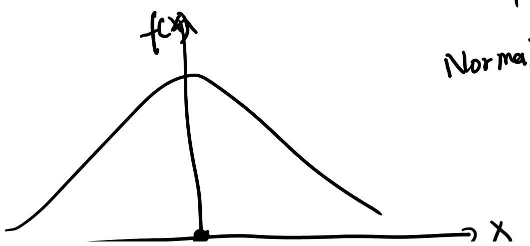
$= 5050$

Euclid's Elements

Regular 17-polygon

3-polygon △

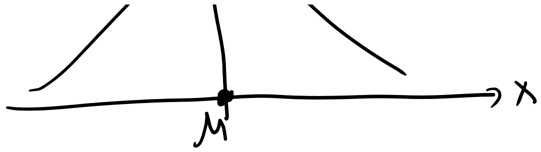
4-polygon □



Normal Distribution

symmetric.

Bell curve



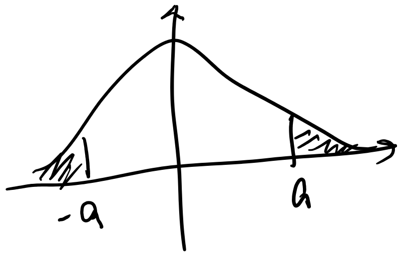
Bell curve

Standard Normal Distribution

$Z \sim N(0, 1)$ Z = standard normal distribution

$p(Z \le -a) = p(Z \ge a)$

$= 1 - p(Z \ge -a)$



$\Phi(a) = p(Z \le a)$

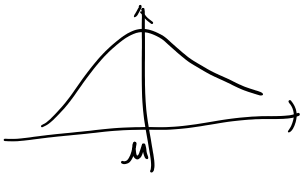
Values of $\Phi(z)$ are given in Z Table.

Central Limit Theorem:

For any R.V., when we repeat random experiment. when the replicates become large enough, the mean of this R.V. tends to have a normal distribution.

$\boxed{\text{mean of } X}$

$\bar{X} \sim N(\mu, \sigma^2)$



Example:

(a) $p(Z \le 2.36) = \Phi(2.36) = 0.9909$

(b) $p(Z \ge 1.46) =$ ① $1 - \underline{p(Z \le 1.46)} = 1 - 0.9279 = 0.0721$

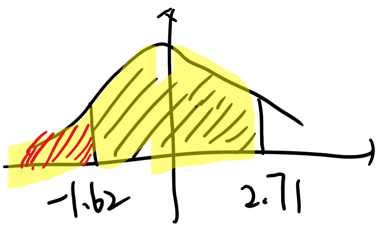
$=$ ② $p(Z \le -1.46) = 0.0721$

symmetric

(c) $p(-1.62 \le Z \le 2.71) = p(Z \le 2.71) - p(Z \le -1.62)$

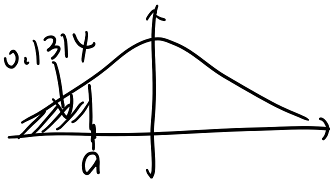
= yellow - red

$= 0.9966 - 0.0526$



(d) $p(Z \le a) = 0.1314$. what's a?

(d) $p(Z \le a) = 0.1314$. what's a ?



$a = -1.12$.

Standize Normal Distribution.

Suppose X follows $N(\mu, \sigma^2)$, then

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

$$\Leftrightarrow \boxed{p(a \le X \le b) = p\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)}$$

Example:

Suppose for some wires has Normal Distribution mean 10 mA and Variance 4 mA².

(a) what's the percentage of time that the wire exceed 13 mA ?

$X \sim N(10, 4)$

$\mu = 10$, $\sigma^2 = 4 \Rightarrow \sigma = 2$

$$p(X \ge 13) = p\left(Z \ge \frac{13-10}{2}\right)$$

$$= p(Z \ge 1.5)$$

$$= p(Z \le -1.5)$$

$$= 0.0668$$

(b) 90% of time, the current exceeds what value ?

$p(X > a) = 0.9$

$\Updownarrow$ $a = ?$

$$p\left(Z > \frac{a-\mu}{\sigma}\right) = 0.9 \Leftrightarrow p\left(Z < \frac{a-\mu}{\sigma}\right) = 0.1$$

$$\Leftrightarrow \frac{a-10}{2} = -1.28$$

$$\Longleftrightarrow \frac{a-10}{2} = -1.28$$

$$a = 7.44$$

(c) If we measure the current using 5 such wires (Independently). What's the probability that the current exceed 13mA in exactly 2 of 5?

① $p = P(X \geq 13) = 0.0668$ ① get the p

$n = 5$

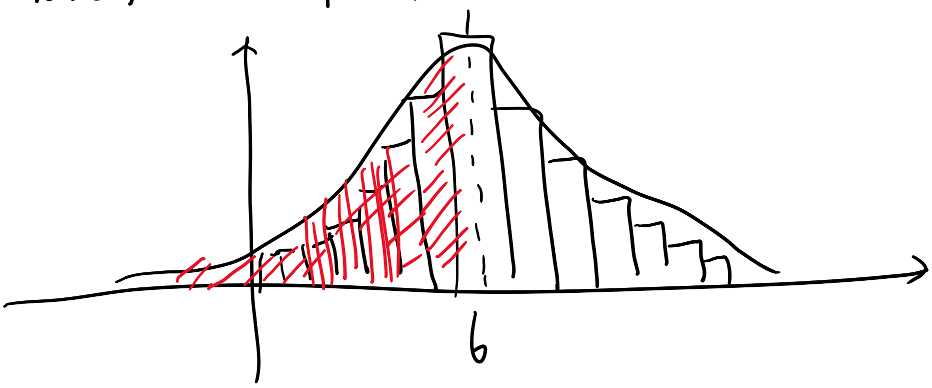
② $\Rightarrow$ Binomial $(n=5,\ p=0.0668)$ ② Find it is Binomial (n, p)

$$P(X=2) = \binom{5}{2} 0.0668^2 (1-0.0668)^3$$

4.7 Normal Approximation to the Binomial.

$E(X) = \mu = np$ $\rangle$ Binomial (n, p)

$Var(X) = \sigma^2 = np(1-p)$



$P(X \leq 6)$

formula:

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

0.5 is called continuity correction.

where np and $n(1-p)$ shall be both greater 5.

Example:

A multiple choice test has 40 questions with 5 choices for each question. If you guess on every question. Find the probability that you get between 3 and 11 inclusive correct.

for each question. ...

the probability that you get between 3 and 11 inclusive correct.

$n=40 \quad p=\frac{1}{5} \Rightarrow \text{Binomial}(40, \frac{1}{5})$

① Exact

$p(3 \le X \le 11)$

$= p(X=3) + p(X=4) + p(X=5) + \dots + p(X=11)$

$= \binom{40}{3}(\frac{1}{5})^3(\frac{4}{5})^{37} + \binom{40}{4}(\frac{1}{5})^4(\frac{4}{5})^{36} + \dots \binom{40}{11}(\frac{1}{5})^{11}(\frac{4}{5})^{29}$

$= 0.9055.$

② Approx

Approximate:

$\mu = np = 40 \cdot \frac{1}{5} = 8$

$\sigma^2 = np(1-p) = \cancel{40}\frac{1}{\cancel{5}} \cdot \frac{4}{5} = \frac{32}{5} =$

$p(3 \le X \le 11)$

$\approx p\left(\frac{3-8-0.5}{\sqrt{32/5}} \le Z \le \frac{11-8+0.5}{\sqrt{32/5}}\right)$

$= p(-2.17 \le Z \le 1.38)$

$= \Phi(1.38) - \Phi(-2.17)$

$= 0.901204$

4.8 Exponential Distribution:

possion: λ = # of mean of success for a certain time interval.

Exponential: $\frac{1}{\lambda}$ = average time of next successs occur.

p.d.f for exponential:

$f(x) = \lambda e^{-\lambda x}, \quad 0 \le x < \infty$

C.D.F for exp:

C.D.F for exp:

$$F(x) = 1 - e^{-\lambda x}, \quad 0 \le x < \infty$$

Mean and Variance:

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

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

Example:

Suppose that the time between arrival of emails at your computer is exp with mean 2 minutes.

(a) what's the probability that you receive an email in a 3 minute period?

$$E(x) = 2 = \frac{1}{\lambda} \qquad \lambda = \frac{1}{2}$$

$$\Rightarrow \begin{cases} f(x) = \frac{1}{2} e^{-\frac{1}{2}x}, & x \geq 0 \\ F(x) = 1 - e^{-\frac{1}{2}x} \end{cases}$$

$$\Rightarrow p(X \le 3) = F(3) = 1 - e^{-\frac{3}{2}}$$

(b) what's the probability that you get an email in the next 3 minutes if you have already been waiting two minutes for an email.

Conditional:

$$p(X \le 5 \mid X > 2) = \frac{p(X < 5 \cap X > 2)}{p(X > 2)}$$

$$= \frac{p(2 < X < 5)}{p(X > 2)}$$

$$= \frac{\quad}{p(X>2)}$$

$$= \frac{F(5) - F(2)}{1 - F(2)}$$

$$= \frac{(1-e^{-\frac{5}{2}}) - (1-e^{-\frac{1}{2}\cdot 2})}{1-(1-e^{-\frac{1}{2}\cdot 2})}$$

$$= 1 - e^{-\frac{3}{2}}$$

memoryless property of exponential:

$$p(X < t_1 + t_2 \mid X > t_1) = p(X < t_2)$$

(c) Determine the length of interval so that the probability that you receive an email in the period is 0.99.

$$p(X < t) = 0.99$$

$$F(t) = 0.99$$

$$1 - e^{-\frac{1}{2}t} = 0.99$$

$$t = -2 \ln \frac{1}{100}$$

Test 1.