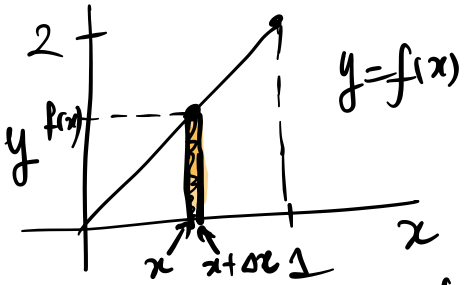
Lecture 34 4/19/2024

X is a continuous r.v. with density $f(x)$

$$\boxed{\begin{array}{l} ① \; f(x) \geq 0 \\ ② \; \int_{-\infty}^{\infty} f(x)\,dx = 1 \end{array}}$$



$$f(x) = \begin{cases} 2x, & 0<x<1 \\ 0 & o/w \end{cases}$$

Area of shaded rectangle

$\approx f(x)\Delta x$

$P(x < X < x+\Delta x)$

$=$ shaded area $\approx f(x)\Delta x$

$$f(x) \approx \frac{P(x < X < x+\Delta x)}{\Delta x}$$

histograms from Data8

$$\underline{\text{height of a bin}} = \frac{\text{\% of data in bin}}{\text{width of bin}}$$

for density histograms
y-axis is density.

$f(x) \to F(x) = P(X \leq x)$

area under $y=f(x)$ over the specified interval: $(-\infty, x]$

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

$$F'(x) = f(x)$$

Discrete r.v.

X, p.m.f $f(x) = P(X=x)$

$F(x) = P(X \leq x)$

$$E(X) = \sum_x x \cdot P(X=x)$$

$$E(X) = \sum_x x \cdot f(x)$$

$$E(g(X)) = \sum_x g(x) P(X=x)$$

$$= \sum_x g(x) f(x)$$

ex: $g(X) = X^2$

$$E(g(X)) = \sum_x x^2 f(x)$$

$$Var(X) = E\left[(X-\mu)^2\right]$$

$$g(X) = (X-\mu)^2$$

$$Var(X) = \sum_x (x-\mu)^2 f(x)$$

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

Show that $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$

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

$$Var(X) = \mathbb{E}(X^2) - \mu^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

All the properties of expectation & variance carry over. a, b etc are always constants

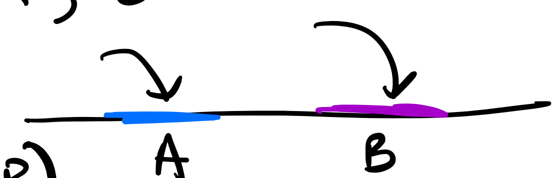
① $\mathbb{E}(aX+b) = a\mathbb{E}(X) + b$

② $Var(aX+b) = a^2 Var(X)$

③ $SD(aX+b) = |a|\, SD(X)$

④ $\mathbb{E}(aX+bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$

⑤ We say that X, Y are independent if for every interval A, B



$$P(X \in A, Y \in B) = P(X \in A)\,P(Y \in B)$$

If X, Y are independent, then

$$Var(X+Y) = Var(X) + Var(Y)$$

$$Var(X-Y) = Var(X) + Var(Y)$$

TWO SPECIAL DISTRIBUTIONS

① exponential dsn

② normal dsn

---

① Exponential dsn.

We say that X has the exponential dsn if its density function $f(x) = \begin{cases} e^{-x}, & x>0 \\ 0, & o/w. \end{cases}$

This particular dsn $(f(x) = e^{-x}, x>0)$ is called the exp. dsn with rate 1.

X is called an exponential r.v. with rate 1

$$X \sim exp(1)$$

In general, we say that $X \sim exp(\lambda)$ ("X has the exponential dsn with rate λ")

if $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x>0, \lambda>0 \\ 0, & o/w. \end{cases}$

λ is some positive constant.

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{0}^{x} \lambda e^{-\lambda t}\,dt = \lambda \int_{0}^{x} e^{-\lambda t}\,dt$$

$$= \cancel{\lambda}\left[-\frac{1}{\cancel{\lambda}} e^{-\lambda t}\Big|_{0}^{x}\right] = 1-e^{-\lambda x}$$

$$F(x) = \begin{cases} 1-e^{-\lambda x} & x>0,\ \lambda>0 \\ 0, & x \leq 0 \end{cases}$$

We usually use T to denote an exponential r.v.

$$T \sim \exp(\lambda) \Rightarrow P(T \leq x) = 1-e^{-\lambda x},\ x,\lambda>0$$

$$P(T>x) = \cancel{1}-(\cancel{1}+e^{-\lambda x}) = e^{-\lambda x}$$

MEMORYLESS PROPERTY OF $T \sim \exp(\lambda)$

$$P(T>s+t \mid T>t) = P(T>s)$$

$$\mathbb{E}(T) = \frac{1}{\lambda} \leftarrow \text{To obtain this, integrate}$$

$$\int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_{0}^{\infty} x \lambda e^{-\lambda x}\,dx$$

$$\mathrm{Var}(T) = \frac{1}{\lambda^2},\ SD(T) = \mathbb{E}(T) = \frac{1}{\lambda}$$

$P(T>t)$ is the prob. that T "survives" past time t.

So we define the survival function $S(t)$ by $S(t) = P(T>t)$

If T models the lifetime of an object $S(t)$ is the chance that it last past time t.

The "memoryless" property of the exp. dsn just means that object "forgets" it has lasted for time t.

$$P(T>s+t \mid T>t) = P(T>s)$$

Median of the exponential distribution.

this is the value s.t. half the dsn is to left of it & half to the right.

If m is the median of the exp. dsn,

$$\underbrace{P(T<m)}_{F(m)} = \underbrace{P(T>m)}_{1-F(m)}$$

$T \sim \exp(\lambda)$

$F(t) = 1 - e^{-\lambda t}$

$$1 - e^{-\lambda h} = \underbrace{e^{-\lambda h}}_{S(h) = P(T > h)}$$

$F(h)$

h for half-life

$$2e^{-\lambda h} = 1$$

$$\boxed{e^{-\lambda h} = \frac{1}{2}} \Rightarrow e^{\lambda h} = 2$$

h is the halfway point of the dsn

$$\lambda h = \ln(2) = \log(2)$$

$$h = \frac{\log(2)}{\lambda} = \log(2) \cdot E(T)$$

log is always ln unless stated otherwise

$$\boxed{h = \underset{\approx 0.69}{\log(2)} \cdot E(T)}$$

$h < E(T)$

Median < average