Lecture 12: Discrete Vs. Continous, the Uniform

discrete	continuous
X	X
PMF $P(X=x)$	PDF $f_X(x) = F_X'(x)$ $[P(X=x)=0]$
CDF $F_X(x) = P(X \leq x)$	CDF $F_X(x) = P(X \leq x)$
$E(X) = \sum_x x P(X=x)$	$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$
$var(X) = E(X^2) - (EX)^2$	$var(X) = E(X^2) - (EX)^2$

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

PDF (probability density function)

A Random variable X has PDF $f(x)$ if

$$P(a \leq X \leq b) = \int_a^b f(x) dx \text{ , for all } a, b.$$

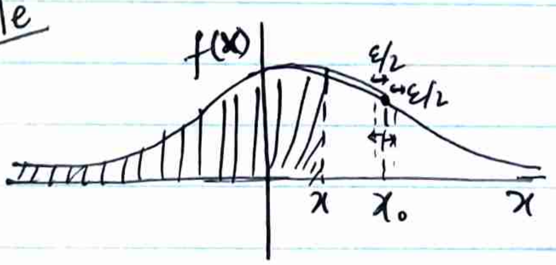
To be valid, $f(x) \geq 0$

and. $\int_{-\infty}^{\infty} f(x) dx = 1$

$a = b \Rightarrow \int_a^a f(x) dx = 0$

Area under the curve from a to a which is 0.

Example



density → probability per unit of length

$f(x_0) \cdot \varepsilon \approx P(X \in (x_0 - \varepsilon/2, x_0 + \varepsilon/2))$

for $\varepsilon > 0$ very small

interval of length ε

By multiplying with ε, we are converting $f(x_0)$ to probability scale instead of density scale

✓ If X has PDF f, the CDF is $F(x) = P(X \leq x)$

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

✓ If X has CDF F (and X is a continuous random variable), then

$f(x) = F'(x)$ (by fundamental theory of Calculus) FTC

$$P(a < X < b) = \int_a^b f(x)\,dx = F(b) - F(a) \quad [\text{by FTC}]$$

Note: Continuous random variable $\longrightarrow$ We have a CDF which has a derivative.

Variance, Var(X)

→ Variance is a measure of how spread out the distribution is.

→ That is, on average, how far is x from ~~the mean~~ its mean?

$$Var(X) = E(X - EX)$$

→ But if we do this, we would always get 0.

By linearity,

$$E(X-EX) = E(X) - E(EX)$$
$$= \cancel{E(X)} - \cancel{E(X)}$$
$$= 0$$

Possible fix –

$$Var(X) = E(|X - EX|)$$

↓

But absolute values are annoying to deal with as they are not differentiable everywhere.

Standard fix:

$$Var(X) = E(X - EX)^2$$

→ We have changed the unit. (became $unit^2$)

Standard deviation : $SD(X) = \sqrt{Var(X)}$

a constant

→ We changed the unit during variance calculation so we changed it back again.

Another way of computing variance

$$Var(X) = E(X^2 - 2X(EX) + (EX)^2)$$
$$= E(X^2) - 2(EX)E(X) + (EX)^2$$
$$= E(X^2) - (EX)^2$$

[(EX) is a constant]

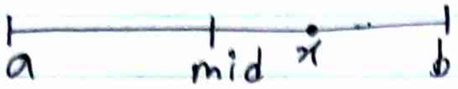
$(EX) = E(X)$

$E(X^2)$ ↓ Squared first, then the average

$(EX)^2$ ↓ Averaged first, the the square

Notation: $EX^2 = E(X^2)$ (Standard Convention)

Uniform distribution, $Unif(a,b)$



"we want to pick a completely random point in $[a,b]$"

The uniform means that probability is proportional to length.

$\boxed{\text{prob.} \propto \text{length}}$

PDF

$$f(x) = \begin{cases} c, & \text{if } a \le x \le b \\ 0, & \text{otherwise} \end{cases} \Rightarrow 1 = \int_a^b c\,dx \Rightarrow c = \frac{1}{(b-a)}$$

CDF

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \overset{x>a}{=} \int_a^x f(t)\,dt \Rightarrow \begin{cases} 0, & \text{if } x < a \\ \frac{x-a}{b-a}, & \text{if } a \le x \le b \\ 1, & \text{if } x > b \end{cases}$$

probability is, as we increase x, the prob. is increasing linearly

$$E(X) = \int_a^b \frac{x}{b-a}\,dx = \frac{x^2}{2(b-a)}\Bigg|_a^b$$

$$= \frac{1}{2(\cancel{b-a})}(\cancel{b-a})(b+a)$$

random variable $= (a+b)/2 \longrightarrow$ midpoint (average is in the middle)

$$E(X^2) = E(Y)$$

$$Y = X^2 = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx$$

(Law of the Unconscious Statician LOTUS)

Law of the Unconscious Statistician (LOTUS)

$$\left[E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx\right]$$

X is a r.v. whose PDF is known

$g(x) \longrightarrow$ function of X

Let $U \sim \text{Unif}(0,1)$. $E(U) = 1/2$, $E(U^2) = \int_0^1 u^2 \underbrace{f_U(u)}_{1} du$ ← LOTUS

$= 1/3$

$\text{Var}(U) = E(U^2) - (EU)^2$

$= 1/3 - 1/4$

$= 1/12$

Note: The uniform distribution is the simplest continuous distribution that we could possibly imagine, because the PDF is a constant on some interval.

→ and we have to have some bounded interval here

(we cannot define a uniform distribution on the entire real line)

✓ Uniform is universal

→ Universality of the uniform means that given a uniform, we can create any distribution that we want.

Let $U \sim \text{Unif}(0,1)$, F be a CDF (Assume F is strictly increasing and continuous function → will have inverse)

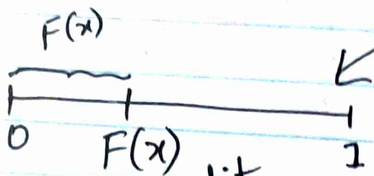
Then Let $X = F^{-1}(U)$. Then $\underbrace{X \sim F}_{X \text{ has CDF } F}$.

What this says is we have CDF we are interested in, we take the inverse CDF, plug in the uniform and then we have constructed a random draw from that distribution we're interested in, F.

<u>proof</u> $P(X \le x) = P(F^{-1}(u) \le x)$

$$= P(u \le F(x)) \quad \left[\begin{array}{c} F^{-1}(u) \le x \\ FF^{-1}(u) \le F(x) \\ u \le F(x) \end{array} \right]$$

$$= F(x)$$



Proportionality constant, here, is 1 as length of the interval is 1.

prob. $\propto$ length (for uniform)