Change of Variable Result for Expectations

X $\qquad \boxed{Y = g(X)}$ } discrete random variables

X discrete

$$E[g(X)] = \sum g(x)\, p_X(x)$$

$$E[Y] = \sum_y y\, p_Y(y)$$

The Case when X and Y are continuous random variables:

$X \in (a, b)$ w.p. 1

$Y = g(X)$ with ~~the~~ setting being identical to the first (basic) COVT theorem. So $Y \in (c, d)$ w.p 1 where $(c, d) = (g(a), g(b))$ if $g \uparrow$

$= (g(b), g(a))$ if $g \downarrow$

$Y = g(X)$

$E[g(X)] = \int_{(a,b)} g(x) f_X(x)\, dx$ (*)

$E[Y] = \int_{(g(a), g(b))} y\, f_Y(y)\, dy$ $\quad g \uparrow$

$f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$

$E(Y) = \int_{(g(a), g(b))} y\, f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)\, dy$

Set $w = g^{-1}(y) \iff y = g(w)$

Then $w \in (a, b)$

$dw = \frac{d}{dy} g^{-1}(y) \cdot dy$

so $E(Y) = \int_{(a,b)} g(w) f_X(w)\, dw$ precisely (*).

Example:

Suppose X follows $Exp(\lambda)$.

$Y = \lfloor X \rfloor \rightarrow$ largest integer not exceeding X. $g(x) = \lfloor x \rfloor$!

continuous $\longrightarrow$ discrete transform

Y has a pmf since $Y \in \{0, 1, 2, \dots\}$

$$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}(x > 0)$$

p.m.f of Y.

Based on Problem 3 1- HW(2), you can show ~~that~~ $P[Y = m] = e^{-\lambda m}(1 - e^{-\lambda})$

$m = 0, 1, 2, \dots$

Y is actually geometric w.p $p = (1 - e^{-\lambda})$

$$E[Y] = \sum_{m=0}^{\infty} m e^{-\lambda m}(1 - e^{-\lambda}) = \frac{e^{-\lambda}}{1 - e^{-\lambda}}$$

$$y = g(x) = \lfloor x \rfloor$$

$$E(Y) = \int_0^\infty g(x) f_X(x)\, dx$$

$$= \int_0^\infty \lfloor x \rfloor \lambda e^{-\lambda x}\, dx$$

between 0 and 1, $\lfloor x \rfloor = 0$ on $[0, 1]$

on $[1, 2]$, $\lfloor x \rfloor = 1 \ldots$

$$= \int_1^2 1 \cdot \lambda e^{-\lambda x}\, dx + \int_2^3 2 \cdot \lambda e^{-\lambda x}\, dx$$

m to $m+1$

$\lfloor x \rfloor = m$

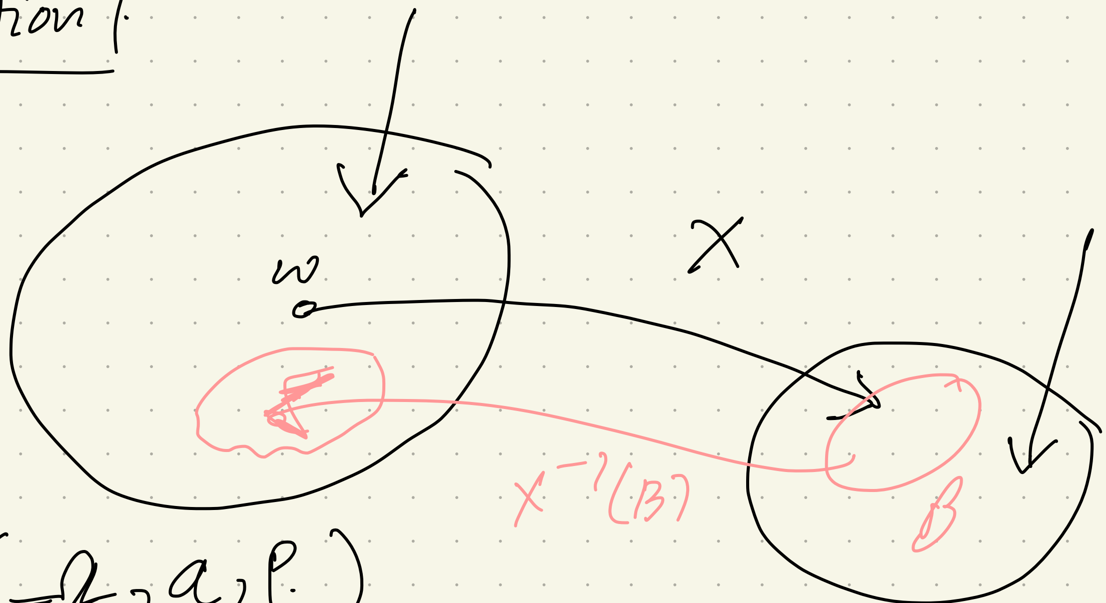
$$+ \ldots$$

$$= \sum_{m=1}^{\infty} m \int_m^{m+1} \lambda e^{-\lambda x}\, dx$$

$$= \sum_{m=1}^{\infty} m \left(e^{-\lambda m} - e^{-\lambda(m+1)}\right)$$

$$= \sum_{m=1}^{\infty} m e^{-\lambda m} \left(1 - e^{-\lambda}\right)$$

Notation:



$(\Omega, \mathcal{A}, P)$ $\quad (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$

P induces a distribution P_X on $\mathcal{B}_{\mathbb{R}}$.

if $B \in \mathcal{B}_{\mathbb{R}}$, $P_X(B) = P(X^{-1}(B))$

Corresponds to a distribution F_n.

which is called the c.d.f of X.

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

$$EX = \int X \cdot dP = \int x \cdot dF_X(x)$$

we interpret as $\sum x_j \, p_X(x_j)$ if X takes values $\{x_1, x_2, \dots\}$

When X is continuous with

p.d.f f_X, then

$$\boxed{\int x \, dF_X(x)} = \int x \, f_X(x) dx$$

When X is mixed:

$$EX = \int x \, dF_X(x)$$

$$= \sum_j x_j \, p_X(x_j) + \int_S x f(x) dx$$

$\{x_1, x_2, \ldots\}$

with $P(X = x_j) > 0$

$\quad\quad\quad\quad\quad\quad \shortparallel$

$\quad\quad\quad\quad p_X(x_j)$

$$\sum_j p_X(x_j) < 1$$

on S, s.t. $S \cap \{x_1, \ldots\} = \phi$,

$P(X \in (a,b)) = \int_a^b f_X(x) dx$, for $(a,b) \subseteq S$

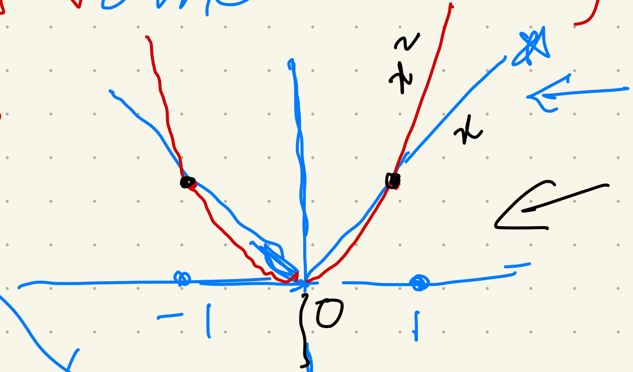
X is a random variable.

Then $\underline{EX}$ minimizes $E[\underline{(X-a)^2}]$ over all a. (M.S.E of X around a).

discrepancy between X and its predictor a.

What about using other kinds of discrepancies?

$\underline{L(x,a)}$: a non-negative function which measures some discrepancy between x and a.



$L(x,a) = \underline{(x-a)^2}$

$\underline{L(x,a)} = |x-a|$

sensitive to large values

What is a^* that minimizes $E(|X-a|)$?

Answer: Median of X i.e

any a^* s.t. $F_X(a^-_*) \le 1/2 \le F_X(a_*)$

minimizes $E(|X-a|)$

What about minimizing $E[(X-a)^2]$

this is EX.

$\boxed{\text{20 kids, 2 giraffes}}$

Mean height gets swayed by giraffes

Median height is not.

anything between 11th and 12th smallest height.

$F(u-) = \lim_{x \uparrow u-} F(x) = P(X < u)$

$F(u) = P(X \le u)$.

$E(|X-a|)$ and let m be a median

so $F(m-) \le 1/2 \le F(m)$

$$E(|X-a|) = \int |x-a| \, dF(x).$$

$$= \int_{(-\infty,a)} (a-x)\, dF(x) + \int_{[a,\infty)} (x-a)\, dF(x)$$

$$= \int_{(-\infty,a)} a \, dF(x) - \int_{(-\infty,a)} x \, dF(x) + \int_{[a,\infty)} x \cdot dF(x)$$

$$- a \int_{[a,\infty)} dF(x)$$

$$= a \, P(X \in (-\infty, a)) - \int_{(-\infty,a)} x \, dF(x)$$

$$\| \quad a F(a-)$$

$$+ \int_{[a,\infty)} x \, dF(x) - a(1 - F(a-))$$

plug in here

$$\mu = \int_{(-\infty,\infty)} x \cdot dF_X(x)$$

$$= \int_{(-\infty,a)} x \, dF_X(x) + \int_{[a,\infty)} x \cdot dF_X(x)$$

$$\int_{(-\infty,a)} x \, dF_X(x) = \mu - \int_{[a,\infty)} x \, dF_X(x) \quad (*)$$

Then:

$$E(|X-a|) = a - 2a(1 - F(a-)) - \mu + 2\int_{[a,\infty)} x \, dF(x)$$

$$= a + 2\int_{[a,\infty)} (x-a) \, dF(x) - \mu$$

holds when $a=m$

Focus on $a \leq m$ } $\underline{m \geq a}$

$$E[|X-a|] - E[|X-m|] \geq 0$$

$$= \underline{(a-m)} + 2\int_{[a,\infty)} (x-a)\, dF(x)$$

$$- 2\int_{[m,\infty)} (x-m)\, dF(x)$$

$$=$$

$$\underline{2\int_{[a,m)} (x-a)\, dF(x)} + 2\int_{[m,\infty)} (x-a)\, dF(x)$$

(via one or two manipulations)

$$= 2(m-a)\left[(1 - \underline{F(m-)}) - \frac{1}{2}\right]$$

$$+ \quad 2\int_{[a,m)} (x-a)\, dF(x)$$

≥ 0

≥ 0

$\frac{1}{2} - F(m-) \geq 0$

Notion of variability.

<u>Mean</u> <u>median</u>

measures of centrality

they're notions of capturing values AROUND the random variable is dispersed!

<u>Critical Notion:</u> How much dispersion does a random variable possess?

Need notions of dispersion.

Variance ← central notion of dispersion

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

mean squared error about average

$$MSE(X, a) = E\left[(X-a)^2\right]$$

$$\left. s.d(X) = \sqrt{Var(X)} \right\}$$

M.A.D. Mean Absolute Deviation

is given $E\left[|X - EX|\right]$.

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

Expand $E\left[(X - EX)^2\right]$ and

then manipulate...

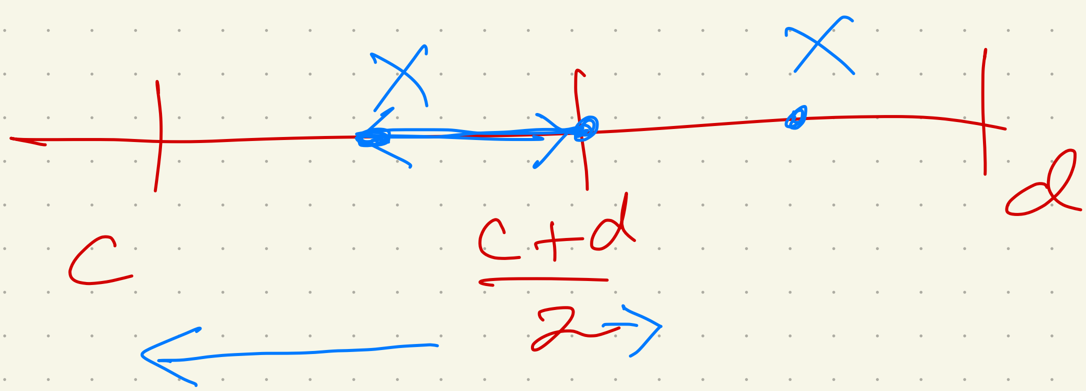
$$Var(aX + b) = Var(aX) = a^2 Var(X)$$

Relation between variance and range.

Ex 3.7: If X is a random variable taking values in $[c, d]$, then

$$\mathrm{Var}(X) \le \frac{(d-c)^2}{4}.$$

$$\left| X - \frac{c+d}{2} \right| \le \frac{d-c}{2}$$



$$E\left[\left(X - \frac{c+d}{2}\right)^2\right] \le \left(\frac{d-c}{2}\right)^2$$

$$E\left[(X - EX)^2\right] \le E\left[\left(X - \frac{c+d}{2}\right)^2\right]$$

a

$X \in [c, d]$ w.p. 1.

Then $P(|X - EX| \geq t)$

$$\leq 2 \exp\left[\frac{-2t^2}{(d-c)^2}\right]$$

Hoeffding's inequality

simple form of concentration inequality . .