- Properties of variance (X = r.v)

  * for any X, $Var(X) \geq 0$
  * If X is bounded, then $Var(X)$ is also bounded.
    ↳ μ is bounded.

Thm 4.3.3 $Var(X) = 0$ iff $\exists$ constant c s.t.

$$P(X = c) = 1$$

pf) $\Leftarrow$ $E(X) = \mu \cdot \Rightarrow c$

$$P((X - \underbrace{c}_{=\mu})^2 = 0) = 1$$

$$Var(X) = E((X - \mu)^2) = 0$$

$\Rightarrow$

$\boxed{\begin{array}{l} \text{Lem } i) P(X \geq a) = 1 \text{ or } P(X \leq a) = 1 \quad (a = \text{constant}) \\ ii) E(X) = a \\ \Rightarrow P(X = a) = 1 \end{array}}$

Since $Var(X) = 0$, $E((X - \mu)^2) = 0$, and

$$P((X - \mu)^2 \geq \underbrace{0}_{a}) = 1$$

$\Rightarrow$ by the Lemma, $P((X - \mu)^2 = 0) = 1$

$$\Leftrightarrow P(X = \mu) = 1$$

by set $c = \mu$, the statement holds.

Thm 4-3-4 r.v. $Y = aX + b$ (a, b = constant)

$Var(Y) = a^2 \cdot Var(X)$

pf) $Var(Y) = E(((aX+b) - \underbrace{E(aX+b)}_{a \cdot E(X) + b})^2)$

$= E((a(X-\mu))^2)$

$= E(a^2 (X-\mu)^2) = a^2 \cdot \underbrace{E((X-\mu)^2)}_{Var(X)}$

$= a^2 \cdot Var(X)$

//

Cor $Var(-X) = (-1)^2 \cdot Var(X) = Var(X)$

Thm 4-3-5 $X_1, X_2, \ldots X_n$ = n r.v., Independent

$Var(X_1 + X_2 + \cdots + X_n)$

$= Var(X_1) + Var(X_2) + \cdots + Var(X_n)$

pf) case when $n = 2$,

let $E(X_1) = \mu_1$, $E(X_2) = \mu_2$

$E(X_1 + X_2) = \mu_1 + \mu_2$

$$Var(X_1+X_2) = E\left(\left((X_1+X_2)-(\mu_1+\mu_2)\right)^2\right)$$

$$(X_1-\mu_1)+(X_2-\mu_2)$$

$$= E\Big(\underbrace{(X_1-\mu_1)^2}_{Var(X_1)} + 2\cdot(X_1-\mu_1)(X_2-\mu_2) + \underbrace{(X_2-\mu_2)^2}_{Var(X_2)}\Big)$$

$(\because X_1$ and X_2 are independent$)$

$$= Var(X_1)+Var(X_2)+\cancel{2\cdot E(X_1-\mu_1)\cdot E(X_2-\mu_2)}$$

$$E(X_1)-\mu_1 = 0 \rightarrow 0$$

$$= Var(X_1)+Var(X_2)$$

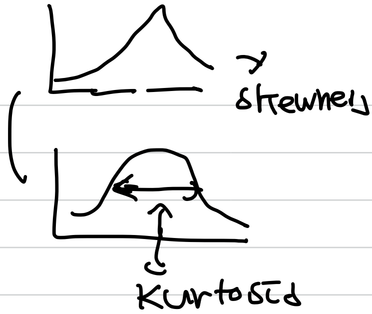
Cor $X_1, X_2, \cdots X_n$ = n independent r.vs

$a_1 \cdots a_n$ = constant.

$$Var(a_1X_1+a_2X_2 \cdots + a_nX_n)$$

$$= a_1^2 Var(X_1)+a_2^2 Var(X_2)+ \cdots + a_n^2 Var(X_n)$$

4-4. Moments



↳ the other measure to summarize the r.v.

Def K-th moment = expectation of X^K $(K \ge 1)$

Def K-th moment exists $\longleftrightarrow$ $E(X^K)$ exists (bounded) $(< \infty)$ $(> -\infty)$

Thm 4-4-1 If K-th moment exists then $\forall j < K$, j-th moment also exists.

pf) We claim that $E(|X|^j)$ has upper bound.

(assume that X is continuous r.v.)

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

LOTUS

$f(x)$: p.d.f. X.

$$= \int_{|x|\le 1} |x|^j f(x)\,dx + \underbrace{\int_{|x|>1} |x|^j f(x)\,dx}_{|x|^j < |x|^k}$$

$$\le \underbrace{\int_{|x|\le 1} f(x)\,dx}_{P(|x|\le 1)\le 1} + \underbrace{\int_{|x|>1} |x|^k f(x)\,dx}_{E(|x|^k)<\infty}$$

$\Rightarrow$ bounded $(<\infty)$

$\therefore$ j-th moment exists

* moment generating function.

Def X = r.v. t = real value

the moment generating function (m.g.f.)

of X, $\psi(t)$ is defined as

psi

$$\psi(t) = E(e^{t\cdot X})$$

- $\psi(t)$ only depends on the distribution of X
- If m.g.f. of two r.vs X and Y are same, they have the same distribution

$(-\varepsilon, \varepsilon)$ ←

Thm X = r.v., $\psi(t)$ = m.g.f of X. where $\psi(t)$ is finite for all values of t some open interval around 0. Then the n-th moment of X, $E(X^n)$ is
$(n \geq 1)$

← n-th derivate.

$$E(X^n) = \psi^{n}(0).$$

Ex) X = r.v. with p.d.f f.

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{otherwise}, \end{cases}$$

m.g.f of $X = E(e^{tX})$

LOTUS → $= \int_0^{\infty} e^{tx} \cdot e^{-x}\,dx = \int_0^{\infty} e^{x(t-1)}\,dx$

$y = (t-1)x$
$dy = (t-1)\,dx$

$t \geq 1 \to \psi(t)$ does not exist

$t < 1$

(when $t<1$) $= \int_0^{-\infty} \frac{1}{t-1} \cdot e^{y} \cdot dy$

$= \frac{1}{t-1} \cdot e^{y} \Big|_0^{-\infty} = \frac{1}{1-t}$

$(-0.5, 0.5)$ → finite.

$$\psi(t) = \frac{1}{1-t} = (1-t)^{-1}$$

$$\text{1-st moment} = \psi'(0) \rightarrow E(X)$$

$$= (1-t)^{-2} = 1$$

$$\text{2nd moment} = \psi''(0) \rightarrow E(X^2)$$

$$= +2(1-t)^{-3} = 2.$$

$$Var(X) = \underline{E(X^2)} - (\underbrace{E(X)}_{1})^2$$

$$= 2 - 1 = 1$$

Thm 4.4.3 : X = r.v. with m.g.f ψ_1

$Y = aX + b$ (a, b = constant)

then m.g.f. of Y, $\underline{\psi_2(t)}$ is

$\forall t$ where $\psi_1(at)$ is finite

$\psi_1(at) = E(e^{atX})$

$$\psi_2(t) = e^{bt} \cdot \psi_1(at)$$

pf) $\psi_2(t) = E(e^{tY}) = E(e^{t(aX+b)})$

$$= e^{tb} \cdot E(e^{taX}) = e^{tb} \cdot \psi_1(at)$$

ex) $\psi(t) = \frac{1}{1-t}$ for $t<1$

r.v. $Y = 3 - 2X$, $\psi_2(t)$ = m.g.f. of Y

Then for $\forall t<1$

$$\psi_2(t) = e^{3t} \cdot \frac{1}{1+2t}$$

Thm 4.4.4 $X_1 \cdots X_n$: n r.vs, independent

$\psi_i(t)$ = m.g.f. of X_i

Let r.v. $Y = X_1 + X_2 \cdots + X_n$.

Then the m.g.f. of Y, $\psi(t)$ is

$\forall t$ where $\psi_i(t)$ is finite $\forall i$.

$$\psi(t) = \prod_{i=1}^{n} \psi_i(t)$$

pf) $\psi(t) = E(e^{tY}) = E(e^{t(X_1+X_2 \cdots + X_n)})$

$= E(e^{tX_1} \cdot e^{tX_2} \cdots e^{t \cdot X_n})$

$\downarrow$ $X_1 \cdots X_n$ are independent

$= E(e^{tX_1}) \cdot E(e^{tX_2}) \cdots E(e^{tX_n})$

$= \prod_{i=1}^{n} \psi_i(t)$.

4-5 The mean and the median

Def A median of distribution of X.
: every number m s.t.
i) $P(X \geq m) \geq \frac{1}{2}$, and
ii) $P(X \leq m) \geq \frac{1}{2}$

i) Multiple numbers can be medians of the distribution
ii) Every distribution must have at least one median.

ex) X = discrete r.v.

$P(X=1) = 0.1$ $\quad$ $P(X=3) = 0.3$

$P(X=2) = 0.2$ $\quad$ $P(X=4) = 0.4$

→ 3 is the median of the distribution of X ? O

i) $P(X \geq 3) = 0.7 > 0.5$

ii) $P(X \leq 3) = 0.6 > 0.5$

ex) X = conti r.v. with p.d.f f.

$$f(x) = \begin{cases} 4x^3 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

i) $P(X \geq m) \geq \frac{1}{2}$

x^4

$$\Rightarrow \int_m^1 4x^3 \, dx \geq \frac{1}{2}$$

$$1 - m^4 \geq \frac{1}{2}$$

$$-m^4 \geq -\frac{1}{2} \Rightarrow m \leq \left(\frac{1}{2}\right)^{1/4}$$

--- ①

ii) $P(X \leq m) \geq \frac{1}{2}$

$$\Rightarrow \int_0^m 4x^3 \, dx \geq \frac{1}{2}$$

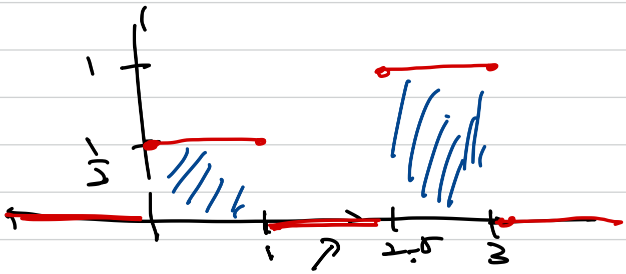
$$m^4 \geq \frac{1}{2}$$

$$m \geq \left(\frac{1}{2}\right)^{1/4} \quad \text{-- ②}$$

$\Rightarrow$ the distribution of X has a unique median $m = \left(\frac{1}{2}\right)^{1/4}$.

ex) X = conti r.v. p.d.f. f.

$$f(x) = \begin{cases} \frac{1}{2} & 0 \le x \le 1 \\ 1 & 2.5 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$



$\forall m$ where $\underline{1 \le m \le 2.5}$ are medians of the distribution of X.

$$\boxed{\begin{aligned} &\text{i) } P(X \ge m) = \tfrac{1}{2} \\ &\text{ii) } P(X \le m) = \tfrac{1}{2} \end{aligned}}$$

//

평균 제곱 오차

— Mean-square error (MSE)

random variable X.

(the value of X is observed from the experiment)

* predict the value of X $\boxed{\text{before}}$ the experiment $\longrightarrow$ d (predicted value)

* Error between X and d.

Def.) The mean-square error (MSE) of prediction d, is $E((X-d)^2)$.

☆☆ Thm 4.5.2 r.v X has mean μ and variance σ^2, then $\forall d$.

$$E((X-\mu)^2) \le E((X-d)^2)$$

$\Rightarrow$ when predict d as μ, one can minimize MSE.

$f(d)$

pf) $E(\underbrace{(X-d)^2}_{X^2-2Xd+d^2}) = E(X^2) - 2E(X)d + d^2$

function of d

$2d - 2E(X) = 0$

$\longrightarrow$

when $d = \overset{\mu}{\overset{\shortparallel}{E(X)}}$, $f(d)$ is minimized.

Sec 4-6 Covariance & Correlation

→ when joint distribution of two r.v.s is given, want to summarize how much these two r.v.s depend each other

Def X, Y = r.v.s with finite means μ_X and μ_Y resp. Then the covariance of X and Y ($\mathrm{Cov}(X,Y)$) is defined as

$+$
0
$-$

$$\mathrm{Cov}(X,Y) = E\left(\underline{(X-\mu_X)}\,\underline{(Y-\mu_Y)} \right)$$

* If the variance of X and Y are bounded, $\mathrm{Cov}(X,Y)$ is bounded

ex) X, Y = conti r-vs with joint p.d.f. f.

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

Cov (X, Y)

$$\int_0^1 2xy + \frac{1}{2}\, dy$$

$$xy^2 + \frac{1}{2}y \Big|_0^1$$

$$\mu_X = \int_0^1 x \cdot \underbrace{\boxed{\int_0^1 f(x,y)\, dy}}_{f_1(x)}\, dx$$

$$= \int_0^1 x \cdot \left(x + \frac{1}{2}\right) dx$$

$$= \frac{1}{3}x^3 + \frac{1}{4}x^2 \Big|_0^1 = \frac{7}{12}.$$

$$\mu_Y = \frac{7}{12} \quad (x \text{ and } y \text{ are symmetric})$$

$$\mathrm{Cov}(X,Y) = E\left(\underline{\left(X - \frac{7}{12}\right) \cdot \left(Y - \frac{7}{12}\right)}\right)$$

LOTUS

$$\Longrightarrow \int_0^1 \int_0^1 \left(x - \frac{7}{12}\right) \cdot \left(y - \frac{7}{12}\right) \cdot \left(2xy + \frac{1}{2}\right) dx\, dy$$

$$= \frac{1}{144}.$$

Thm 4.6.1 X, Y = r.v with finite variance σ_X^2 and σ_Y^2 resp.

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$$

pf) Let $\mu_X = E(X)$, $\mu_Y = E(Y)$

$$\mathrm{Cov}(X,Y) = E((X-\mu_X)(Y-\mu_Y))$$

$$= E(\underline{XY} - \underline{\mu_Y X} - \underline{\mu_X Y} + \underline{\mu_X\mu_Y})$$

$$= E(XY) - \mu_Y \underbrace{E(X)}_{\mu_X} - \mu_X \underbrace{E(Y)}_{\mu_Y} + \mu_X\mu_Y$$

$$\underline{\qquad\qquad\qquad\qquad}$$
$$-2\mu_X\mu_Y$$

$$= E(XY) - \underbrace{\mu_X\mu_Y}_{E(X)\cdot E(Y)}$$

//

Def (correlation) → scaled covariance
(make covariance to not driven by arbitrary scaled r.v.s)

the correlation of two r.v.s X and Y,

$\rho(X,Y)$ (Corr(X,Y))

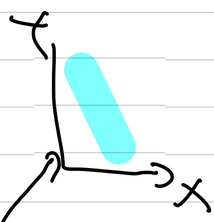
rho

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

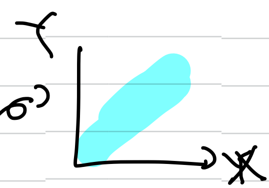
$(\sigma_X^2 = Var(X),\ \sigma_Y^2 = Var(Y))$

Thm 4-6-3 $(Cov(X,Y))^2 \le \sigma_X^2 \cdot \sigma_Y^2$

by Couchy-Schwarz inequality.

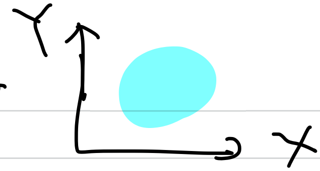


$$\therefore \frac{(Cov(X,Y))^2}{\sigma_X^2 \sigma_Y^2} \le 1$$



$\Rightarrow -1 \le \rho(X,Y) \le 1$

- $\rho(X,Y) > 0$ → positive correlated
- $\rho(X,Y) = 0$ → uncorrelated
- $\rho(X,Y) < 0$ → negative correlated

$\rho(X,Y)=0 \longrightarrow$ 

- Properties of Covariance & correlation

Thm If two r.v.s X and Y are independent, and σ_X^2 and σ_Y^2 are finite, then

$$Cov(X,Y)=0$$

pf) $Cov(X,Y) = \underbrace{E(XY)}_{E(X)\cdot E(Y)} - E(X)E(Y)$

(∵ X and Y are independent) $= 0$.

Converse is not true!!

Thm 4.6.5 X = r.v. with finite σ_X^2

$Y = aX+b$ (a, b : constant)

if $a>0$, then $\rho(X,Y)=1$

$a<0$, " $\rho(X,Y)=-1$.

pf) $\mu_X = E(X)$

$\mu_Y = E(aX+b) = a\mu_X + b$

$Cov(X,Y) = E((X-\mu_X)(Y-\mu_Y))$

$Y - \mu_Y = aX + \cancel{b} - (a\mu_X + \cancel{b})$

$a\underbrace{(X-\mu_X)}$

$= E(\underline{a}(X-\mu_X)^2)$

$= a \cdot \sigma_X^2$

$\sigma_Y^2 = a^2 \cdot \sigma_X^2$

$\sigma_Y = |a| \cdot \sigma_X$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{a \cdot \cancel{\sigma_X^2}}{\cancel{\sigma_X} \cdot |a| \cdot \cancel{\sigma_X}}$$

$$= \frac{a}{|a|}$$

$\therefore$ $a > 0 \rightarrow \frac{a}{a} = 1$

$a < 0 \rightarrow \frac{a}{-a} = -1.$

Thm 4-6-6 X, Y = r-vs with finite σ_X^2 and σ_Y^2

$$Var(X+Y) = Var(X) + Var(Y) + 2 \cdot Cov(X, Y)$$

pf) $Var(X+Y) = E\left(\left(\underbrace{(X+Y) - (\mu_X + \mu_Y)}_{(X-\mu_X)+(Y-\mu_Y)} \right)^2 \right)$

$$= E\Big(\underbrace{(X-\mu_X)^2}_{Var(X)} + \underbrace{2(X-\mu_X)(Y-\mu_Y)}_{Cov(X,Y)} + \underbrace{(Y-\mu_Y)^2}_{Var(Y)} \Big)$$

$$= Var(X) + 2\,Cov(X,Y) + Var(Y)_{,,}$$

Cor: X, Y = r-v.
a, b, c = constant

$$Var(aX + bY + \underline{c}) = Var(aX + bY)$$

$$= a^2 Var(X) + b^2 Var(Y) + 2\,\underbrace{Cov(aX, bY)}_{2ab\,Cov(X,Y)}$$

check! → ($2ab\,Cov(X,Y)$

Thm $X_1 \cdots X_n$ = r.v.s with finite $Var(X_1), Var(X_2), \cdots, Var(X_3)$. Then

$$Var\left(\sum_{i=1}^{n} X_i\right)$$

- $Cov(X_1, X_2)$
- ~~$Cov(X_2, X_1)$~~

$$= \sum_{i=1}^{n} Var(X_i) + 2\sum\sum_{i<j} Cov(X_i, X_j)$$

pf) $Cov(X, X)$

$$= E\left((X-\mu_X)^2\right) = Var(X)$$

$$Var\left(\sum_{i=1}^{n} X_i\right) = Cov\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right)$$

check!

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(X_i, X_j)$$

$$= \underbrace{\sum\sum_{i=j} Cov(X_i, X_j)}_{\sum_{i=1}^{n} Var(X_i)} + \underbrace{\sum\sum_{i\neq j} Cov(X_i, X_j)}_{2\sum\sum_{i<j} Cov(X_i, X_j)}$$

$Cov(X, Y) = Cov(Y, X)$

$\downarrow$

$$\sum_{i=1}^{n} Var(X_i) + 2\sum\sum_{i<j} Cov(X_i, X_j)$$