10/18

Last time:

$$\boxed{Var(X) = E\left[(X - E(X))^2\right]}$$

For discrete case:

$$Var(X) = \sum_i (x_i - \mu)^2 p(x_i)$$

where, $\mu = E(X)$.

For continuous case:

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

Theorem: $\boxed{Var(X) = E(X^2) - (E(X))^2}$

Proof: $Var(X) = E(X-\mu)^2$ (with $\mu = E(X)$)

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

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

$$= E(X^2) + \mu^2 - 2\mu\cdot\mu$$

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

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

$$E(\mu^2) = \mu^2 E(1) = \mu^2$$

Example: Find $E(X)$ and $Var(X)$ if $X \sim Unif([0,1])$

Solution: pdf $f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$

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

$= \int_0^1 x \cdot 1\, dx = \boxed{\frac{1}{2}}$ //

$\boxed{Var(X)} = E(X^2) - (E(X))^2$

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

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

$= \int_0^1 x^2 \cdot 1\, dx$

$= \frac{1}{3}$

$= \frac{1}{3} - \frac{1}{4}$

$= \boxed{\frac{1}{12}}$

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

Example: Normal distribution $X \sim N(\mu, \sigma^2)$

Std. deviation $= \sqrt{\sigma^2} = \sigma$ //

Chebyshev's inequality

Let X be a random variable with

$\boxed{\text{mean}} = \mu$, $\boxed{\text{variance}} = \sigma^2$

$E(X)$ $\quad$ $Var(X)$

Then, for any $t > 0$

$$\boxed{P(|X - \mu| > t) \le \frac{\sigma^2}{t^2}}$$

, σ ~~fin~~ is finite

Remark:

$\sigma = \frac{1}{10}$, $\mu = 0$

$$P(|X - 0| > t) \le \frac{1}{100t^2}$$

If $t = 1$

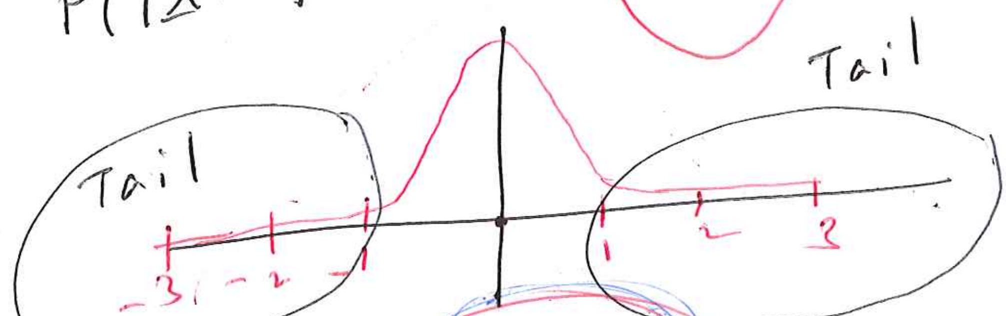
$$P(|X - 0| > 1) \le \frac{1}{100}$$

If $t = 2$

$$P(|X - 0| > 2) \le \frac{1}{400}$$

If $t = 3$

$$P(|X - 0| > 3) \le \frac{1}{900} \ldots$$



This is NOT true if $\sigma = +\infty$

"Fat-tail distribution"

Proof:

For continuous case:

$$\boxed{\sigma^2} = \int_{-\infty}^{\infty} \overset{\geq 0}{(x-\mu)^2}\, \overset{\geq 0}{f(x)}\, dx$$

$$\geq \int_A (x-\mu)^2 f(x)\, dx$$

$$\geq \int_A t^2 f(x)\, dx$$

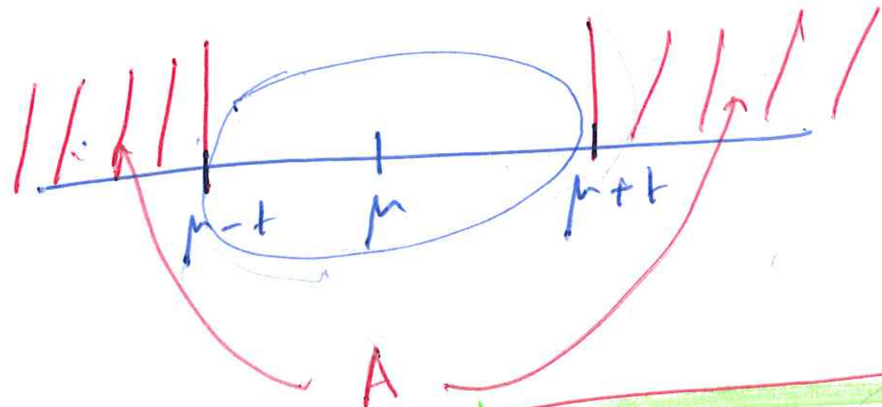
$$= \boxed{t^2 \int_A f(x)\, dx}$$

$$\boxed{A = \{x : |x-\mu| > t\}}$$

$|x-\mu| > t$

$x-\mu > t$ or $x-\mu < -t$

$x > \mu + t$ or $x < \mu - t$



On A:

$|x-\mu| > t$

$\Rightarrow (x-\mu)^2 > t^2$

$$\boxed{P(A) = \int_A f(x)\, dx}$$

Hence

$$\sigma^2 \geq t^2 P(A)$$

So, $$P(A) \leq \frac{\sigma^2}{t^2}$$

$$P(|x-\mu| > t) \leq \frac{\sigma^2}{t^2} \quad //$$

Corollary: If $Var(X) = 0$
Then $P(X = \mu) = 1$ ✓

Proof: $\boxed{\text{Suppose } P(X = \mu) < 1}$ ~~by~~

Then for some $t > 0$ ~~Chebyshev's inequality)~~ we have

$$P(|X - \mu| > t) > 0$$

However, by Chebyshev's inequality we have $P(|X - \mu| > t) = 0$ for all t.

Contradiction!

Hence, $\boxed{P(X = \mu) = 1}$ //

Question: Does this mean $\boxed{X = \mu}$? **NO!** (in general)

For $\boxed{\text{continuous random variable}}$ ($Var(X) \neq 0$)

$$P(X = \mu) = 0$$

✓ $\boxed{P(\omega \mid X(\omega) = \mu) = 0}$

$\mathbb{R}$: set of real numbers

X : we pick up a ~~rational~~ irrational number ~~($\mathbb{R}$)~~ from $\mathbb{R}$.
(numbers like π, e, $\sqrt{2}$, ...)

$P(X) = 1$ // ~~(a.s.)~~

It does NOT mean the entire $\mathbb{R}$ is irrational
(Of course there are rational numbers, like $0, \frac{1}{2}, \ldots$
$\frac{1}{2}, \frac{1}{4}, \ldots$)

Covariance and Correlation

If X and Y are jointly distributed random variables with $\mu_X = E(X)$ and $\mu_Y = E(Y)$,

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

provided the expectation exists.

Remark: If the joint distribution of X and Y is given by $f(x,y)$

$$\boxed{Cov(X,Y) = \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty}(x-\mu_X)(y-\mu_Y)f(x,y)\,dx\,dy}$$

Theorem: $Cov(X,Y) = E(XY) - E(X)E(Y)$

Proof:

$$\boxed{Cov(X,Y)} \overset{def}{=} E\big((X-\mu_X)(Y-\mu_Y)\big)$$

$$= E\big(XY - \mu_Y X - \mu_X Y + \mu_X \cdot \mu_Y\big)$$

$$= E(XY) - \mu_Y E(X) - \mu_X \cdot E(Y) + \mu_X \cdot \mu_Y \underbrace{E(1)}_{1}$$

$\mu_X = E(X)$
$\mu_Y = E(Y)$

$$= E(XY) - \mu_Y \cdot \mu_X - \cancel{\mu_X \cdot \mu_Y} + \cancel{\mu_X \cdot \mu_Y}$$

$$= \boxed{E(XY) - E(X)E(Y)}$$

Remark: Suppose $Y = X$

$$\boxed{Cov(X, X)} = E(X^2) - (E(X))^2$$
$$= \boxed{Var(X)}$$

Example:

$$Var(X+Y) = Cov(X+Y, X+Y)$$
$$= E((X+Y)(X+Y)) - E(X+Y)E(X+Y)$$
$$= E(X^2) + 2E(XY) + E(Y^2) - (E(X)^2 + E(Y)^2 + 2E(X)E(Y))$$
$$= \boxed{(E(X^2) - E(X)^2)} + \boxed{(E(Y^2) - E(Y)^2)} + 2\boxed{(E(XY) - E(X)E(Y))}$$

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

Remark: Suppose $X \perp Y$

Then $E(XY) = E(X)E(Y)$

Hence, $Cov(X,Y) = E(XY) - E(X)E(Y)$
$= 0$

Corollary: If $X \perp Y$,

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

[Please note: $E(X+Y) = E(X) + E(Y)$ even when X and Y are NOT independent]

Correlation coefficient

If X and Y are jointly distributed random variables and $Var(X)$, $Var(Y)$ exist and non-zero, then the correlation coefficient of X and Y denoted by ρ is:

$$\boxed{\rho = \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}}}$$

Theorem: $\boxed{-1 \le \rho \le +1}$

If $\rho = \pm 1$ then $\boxed{P(Y = a + bX) = 1}$ for some constants a and b.

$(Y = a + bX)$

Proof:

$$0 \le \boxed{\frac{1}{2} \underbrace{Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right)}_{\sigma^2}}$$

~~$E(X \;)^2$~~

$$= \frac{1}{2}\left(Var\left(\frac{X}{\sigma_X}\right) + Var\left(\frac{Y}{\sigma_Y}\right) + 2\,cov\left(\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}\right)\right)$$

$$\boxed{Var(a + \lambda X) = \lambda^2 Var(X)}$$

$$= \frac{1}{2}\cdot\left(\frac{1}{\cancel{\sigma_X^2}} \cancel{Var(X)} + \frac{1}{\cancel{\sigma_Y^2}} \cancel{Var(Y)} + \frac{2}{\sigma_X \sigma_Y} Cov(X, Y)\right)$$

$$= \frac{1}{2}(1 + 1 + 2\rho)$$

$$= 1 + \rho$$

$$\boxed{Cov\left(\frac{X}{\sigma_x}, \frac{Y}{\sigma_Y}\right) = E\left(\frac{XY}{\sigma_x \sigma_Y}\right) - E\left(\frac{X}{\sigma_x}\right)E\left(\frac{Y}{\sigma_Y}\right) = \frac{1}{\sigma_x \sigma_Y}\left(E(XY) - E(X)E(Y)\right) = \frac{1}{\sigma_X \sigma_Y} Cov(X, Y)}$$

Hence

$$0 \leq 1 + \rho$$

$$\Rightarrow \boxed{\rho \geq -1}$$

Similarly, if we started off with

$$0 \leq \frac{1}{2} Var\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right)$$

we get, $0 \leq 1 - \rho$

$$\Rightarrow \cancel{\rho \geq 1} \quad \boxed{\rho \leq 1}$$

Hence, $\boxed{-1 \leq \rho \leq 1}$

If, $\underline{\underline{\rho = -1}}$ then

$$\frac{1}{2} Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 0$$

$$\Rightarrow Var\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right) = 0.$$

$\Rightarrow$ By the previous corollary

$$P\left(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y} = \mu\right) = 1$$

for some μ.

$$\Rightarrow P(Y = a + bX) = 1$$ for some a and b //

Similarly for $\rho = +1$

$$\frac{1}{2} \operatorname{Var}\left(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right) = 0$$

$$\Rightarrow P(Y = a + bX) = 1$$

for some a and b. //