Note 23

$\boxed{2}$ Independent Sum

- Linear combination $a_1 X_1 + \cdots + a_n X_n$ of indep. RVs $X_1, \cdots, X_n$ are particularly important.

THM (Properties of Independent Sum) Let

▷ $X_1, \cdots, X_n$ be indep. RVs,

▷ $a_1, \cdots, a_n$ be constants.

Then the linear combination

$$Y := \sum_{i=1}^{n} a_i X_i$$

satisfies

(1) $E(Y) = \sum_{i=1}^{n} a_i E(X_i)$

(2) $Var(Y) = \sum_{i=1}^{n} a_i^2 Var(X_i)$

(3) $M_Y(t) = M_{X_1}(a_1 t) \cdots M_{X_n}(a_n t)$.

Pf) (1) is nothing but the linearity of $E(\cdot)$.

(2) follows from the linearity of $Cov(\cdot,\cdot)$ in both arguments + independence :

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

$$\overset{\text{(linearity)}}{=} \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \, Cov(X_i, X_j)$$

$$\overset{\text{(indep.)}}{=} \sum_{i=1}^{n} a_i^2 \, Var(X_i) \longleftarrow \because X_1, \cdots, X_n \text{ indep} \Rightarrow Cov(X_i, X_j) = \begin{cases} Var(X_i), & j = i \\ 0, & j \neq i. \end{cases}$$

(3) follows from :

$$M_Y(t) = E[e^{tY}] = E[e^{t(a_1 X_1 + \cdots + a_n X_n)}] = E[e^{a_1 t X_1} \cdots e^{a_n t X_n}]$$

$$\overset{\text{(indep)}}{=} E[e^{a_1 t X_1}] \cdots E[e^{a_n t X_n}] = M_{X_1}(a_1 t) \cdots M_{X_n}(a_n t).$$

$\square$

COR If $X_1, \cdots, X_n$ are i.i.d., then $Y = X_1 + \cdots + X_n$ satisfies

$$E(Y) = n\,E(X_1), \qquad Var(Y) = n\,Var(X_1), \qquad M_Y(t) = M_{X_1}(t)^n.$$

(1) Binomial Distribution Revisited

- Let $X_1, \cdots, X_n$: i.i.d. $\sim$ Bernoulli (p). Then
$$Y = X_1 + \cdots + X_n$$
counts the # of successes out of n indep. trials, and so,
$$Y \sim \text{Binomial}(n, p).$$
It is straightforward to compute :
$$E(X_1) = E(X_1^2) = p, \quad Var(X_1) = p(1-p), \quad M_{X_1}(t) = (1-p) + pe^t.$$
So, we get :
  - $E(Y) = n E(X_1) = np$
  - $Var(Y) = n Var(X_1) = np(1-p)$,
  - $M_Y(t) = M_{X_1}(t)^n = (1-p+pe^t)^n$.

- If $X_1, \cdots, X_k$ are indep. and $X_i \sim \text{Binomial}(n_i, p)$ for each $i = 1, \cdots, k$, then $Y = X_1 + \cdots + X_k$ satisfies
$$M_Y(t) = M_{X_1}(t) \cdots M_{X_k}(t) = (1-p+pe^t)^{n_1 + \cdots + n_k}$$
$$\Rightarrow Y \sim \text{Binomial}(n_1 + \cdots + n_k, p).$$

(2) Negative Binomial Distribution Revisited

- Let $X_1, \cdots, X_n$: i.i.d. $\sim$ Geometric (p). Then we know :
$$E(X_1) = \frac{1}{p}, \quad Var(X_1) = \frac{1-p}{p^2}, \quad M_{X_1}(t) = \frac{pe^t}{1-(1-p)e^t}.$$
So, if $Y = X_1 + \cdots + X_n$, then
$$E(Y) = \frac{n}{p}, \quad Var(Y) = n \cdot \frac{1-p}{p^2}, \quad M_Y(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^n.$$
The MGF of Y shows that $Y \sim$ Negative Binomial (n, p).

(3) Chi-Square Distribution Revisited

▷ If $Z \sim \mathcal{N}(0,1)$, then $X = Z^2$ satisfies:

$$F_X(x) = P(Z^2 \le x) = \begin{cases} P(-\sqrt{x} \le Z \le \sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x}), & \text{if } x > 0, \\ 0, & \text{if } x \le 0. \end{cases}$$

$$\Rightarrow \quad f_X(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, & x > 0 \\ 0, & x \le 0. \end{cases}$$

$$\Rightarrow \quad X \sim \text{Gamma}(\alpha = \tfrac{1}{2}, \theta = 2) = \chi^2(1).$$

▷ It can be proved that $X \sim \text{Gamma}(\alpha, \theta) \Leftrightarrow M_X(t) = (1-\theta t)^{-\alpha}$.

▷ So, if $Z_1, \cdots, Z_n$: i.i.d. $\sim \mathcal{N}(0,1)$, then $X = Z_1^2 + \cdots + Z_n^2$ satisfies:

$$M_X(t) = M_{Z_1^2}(t)^n = \left[(1-2t)^{-1/2}\right]^n = (1-2t)^{-n/2}$$

$$\Rightarrow \quad X \sim \text{Gamma}(\alpha = \tfrac{n}{2}, 2) = \chi^2(n).$$

(4) Sample Mean

Ex If $X_1, \cdots, X_n$: i.i.d. with mean μ & variance σ^2, then

$$\overline{X} = \frac{X_1 + \cdots + X_n}{n} \qquad \text{(sample mean)}$$

Then

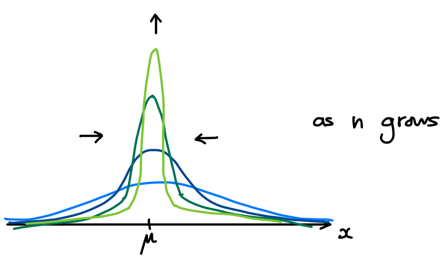
$$\mu_{\overline{X}} = E(\overline{X}) = \frac{1}{n}(\overbrace{\mu + \cdots + \mu}^{n \text{ terms}}) = \mu$$

and

$$\sigma^2_{\overline{X}} = \text{Var}(\overline{X}) = \frac{1}{n^2}(\overbrace{\sigma^2 + \cdots + \sigma^2}^{n \text{ terms}}) = \frac{1}{n}\sigma^2.$$

Note that $\sigma^2_{\overline{X}} \to 0$ as $n \to \infty$, and so, $\overline{X}$ concentrates near μ as n grows. ▯

PDF of $\overline{X}$:



$\boxed{3}$ Common Distribution Revisited

① Distributions arising as an i.i.d. sum

Ex (Bernoulli / Binomial)

- ▷ If $X_1, \cdots, X_n$: i.i.d. $\sim$ Bernoulli(p), then $X_1 + \cdots + X_n \sim$ Binomial(n, p).
- ▷ If $X_1, \cdots, X_k$: indep, $X_i \sim$ Binomial(n_i, p), then $X_1 + \cdots + X_k \sim$ Binomial$(n_1 + \cdots + n_k, p)$

Ex (Geometric / Negative Binomial)

- ▷ If $X_1, \cdots, X_n$: i.i.d. $\sim$ Geometric(p), then $X_1 + \cdots + X_n \sim$ NegativeBinomial(r, p).
- ▷ If $X_1, \cdots, X_k$: indep, $X_i \sim$ NegativeBinomial(r_i, p), then $X_1 + \cdots + X_k \sim$ NegativeBinomial$(r_1 + \cdots + r_k, p)$.

Ex (Exponential / Gamma)

- ▷ If $X_1, \cdots, X_n$: i.i.d. $\sim$ Exponential(θ), then $X_1 + \cdots + X_n \sim$ Gamma(α, θ).
- ▷ If $X_1, \cdots, X_k$: indep, $X_i \sim$ Gamma(α_i, θ), then $X_1 + \cdots + X_k \sim$ Gamma$(\alpha_1 + \cdots + \alpha_k, \theta)$

Ex (Normal^2 / Chi-Square)

- ▷ If $Z_1, \cdots, Z_n$: i.i.d. $\sim \mathcal{N}(0,1)$, then $Z_1^2 + \cdots + Z_n^2 \sim \chi^2(n)$.
- ▷ If $X_1, \cdots, X_k$: indep, $X_i \sim \chi^2(n_i)$, then $X_1 + \cdots + X_k \sim \chi^2(n_1 + \cdots + n_k)$.

② "Infinitely Divisible" distributions

Ex (Poisson) If $X_1, \cdots, X_n$: indep., $X_i \sim$ Poisson(λ_i),

$\Rightarrow X_1 + \cdots + X_n \sim$ Poisson$(\lambda_1 + \cdots + \lambda_n)$

Ex (Normal) If $X_1, \cdots, X_n$: indep., $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$

$\Rightarrow X_1 + \cdots + X_n \sim \mathcal{N}(\mu_1 + \cdots + \mu_n, \sigma_1^2 + \cdots + \sigma_n^2)$.