Note 22

Section 5.3-4. Several Independent Random Variables

I Independence of Several RVs

- Joint PMF/PDF and expectation all extends to n-variables case.
  - ▷ Joint PMF : $p_{X_1,\cdots,X_n}(x_1,\cdots,x_n) = P(X_1 = x_1, \cdots, X_n = x_n)$,
    $P((X_1,\cdots,X_n)\in A) = \sum_{(x_1,\cdots,x_n)\in A} p_{X_1,\cdots,X_n}(x_1,\cdots,x_n)$.
  - ▷ Joint PDF : $P((X_1,\cdots,X_n)\in A) = \int\cdots\int_A f_{X_1,\cdots,X_n}(x_1,\cdots,x_n)\,dx_1\cdots dx_n$,
  - ▷ Expectation : $E[u(X_1,\cdots,X_n)] = \begin{cases} \sum_{x_1,\cdots,x_n} u(x_1,\cdots,x_n)\, p_{X_1,\cdots,X_n}(x_1,\cdots,x_n) \\ \int\cdots\int u(x_1,\cdots,x_n)\, f_{X_1,\cdots,X_n}(x_1,\cdots,x_n)\,dx_1\cdots dx_n \end{cases}$

THM (1) If all X_i's are discrete, then

$$X_i\text{'s indep.} \iff p_{X_1,\cdots,X_n}(x_1,\cdots,x_n) = p_{X}(x_1)\cdots p_{X_n}(x_n).$$

(2) If all X_i's are continuous, then

$$X_i\text{'s indep.} \iff f_{X_1,\cdots,X_n}(x_1,\cdots,x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n)$$

- Independence allows to "factor" more general form of probabilities and expectations :

THM If $X_1,\cdots,X_n$ are independent, then

(1) $P(X_1\in A_1,\cdots,X_n\in A_n) = P(X_1\in A_1)\cdots P(X_n\in A_n)$ for any ranges $A_1,\cdots,A_n \subseteq \mathbb{R}$.

(2) $E[u_1(X_1)\cdots u_n(X_n)] = E[u_1(X_1)]\cdots E[u_n(X_n)]$ for any functions $u_1,\cdots,u_n$.

DEF A sequence of RVs $X_1,\cdots,X_n$ is called a random sample of size n from a common distribution if they are independ and identically distributed.

often abbreviated as i.i.d.

"has the distribution"

Ex Let $X_1 \sim \text{Poisson}(\lambda_1 = 2)$ and $X_2 \sim \text{Poisson}(\lambda_2 = 3)$ be independent. Then

(1) $P(X_1 = 3, X_2 = 5) = P(X_1 = 3) P(X_2 = 5)$

$$= \left(\frac{\lambda_1^3}{3!} e^{-\lambda_1}\right)\left(\frac{\lambda_2^5}{5!} e^{-\lambda_2}\right)$$

$$= \left(\frac{8}{6} e^{-2}\right)\left(\frac{243}{120} e^{-3}\right) = \cdots .$$

(2) $P(X_1 + X_2 = 1) = P(X_1 = 0, X_2 = 1) + P(X_1 = 1, X_2 = 0)$

$$= (e^{-\lambda_1})(\lambda_2 e^{-\lambda_2}) + (\lambda_1 e^{-\lambda_1})(e^{-\lambda_2})$$

$$= (\lambda_1 + \lambda_2) e^{-(\lambda_1 + \lambda_2)} .$$

$\square$

Ex If X, Y : indep. and

$$f_X(x) = \begin{cases} \frac{2}{x^3}, & 1 < x < \infty, \\ 0, & \text{else} \end{cases} \qquad f_Y(y) = \begin{cases} \frac{3}{y^4}, & 1 < y < \infty, \\ 0, & \text{else}, \end{cases}$$

Note These types of dist. are called Pareto distributions.

Then

$$P(X < Y) = \iint_{x<y} f_{X,Y}(x,y)\, dx\,dy \overset{\text{(indep)}}{=} \iint_{x<y} f_X(x) f_Y(y)\, dx\,dy$$

$$= \int_1^\infty \left(\int_x^\infty \frac{2}{x^3} \cdot \frac{3}{y^4}\, dy \right) dx = \int_1^\infty \frac{2}{x^3} \cdot \frac{1}{x^3}\, dx = \frac{2}{5} .$$

$\square$

Ex Let X_1, X_2, X_3 be indep. and $X_i \sim \text{Exponential}(\text{rate } \lambda_i)$ for $i = 1, 2, 3$, i.e.,

$$f_{X_i}(x) = \lambda_i e^{-\lambda_i x}, \quad x > 0.$$

If $Y = \min\{X_1, X_2, X_3\}$, then for $x > 0$,

$$P(Y > x) = P(X_1 > x, X_2 > x, X_3 > x)$$

$$= P(X_1 > x) P(X_2 > x) P(X_3 > x)$$

$$= e^{-\lambda_1 x} \cdot e^{-\lambda_2 x} \cdot e^{-\lambda_3 x}$$

$$= e^{-(\lambda_1 + \lambda_2 + \lambda_3) x} .$$

$\Rightarrow Y \sim \text{Exponential}(\lambda_1 + \lambda_2 + \lambda_3)$.

$\square$

Ex If X, Y : i.i.d. with the common PDF f, then

(1) $$P(X<Y) = \iint_{x<y} f(x)f(y)\,dx\,dy \qquad (\because \text{ independence})$$

$$= \iint_{y<x} f(y)f(x)\,dx\,dy \qquad (\because \text{ renaming dummy variables})$$

$$= P(Y<X).$$

(2) $$P(X=Y) = \iint_{x=y} f(x)f(y)\,dx\,dy = 0 \qquad (\because \text{ line has zero area.})$$

(3) So, $$P(X<Y) \overset{(1)}{=} \frac{1}{2}\big(P(X<Y)+P(Y<X)\big) = \frac{1}{2}\big(1-P(X=Y)\big) \overset{(2)}{=} \frac{1}{2}.$$

$\square$

Ex (Beta Distribution) If $X \sim \text{Gamma}(\alpha, 1)$ and $Y \sim \text{Gamma}(\beta, 1)$ are indep., then

$$U = \frac{X}{X+Y}$$

is a RV taking values between 0 and 1. Given $X = x$, the conditional PDF of U is

$$f_{U|X}(u|x) = \frac{d}{du} P\left(\frac{X}{X+Y} \leq u \,\middle|\, X=x\right) = \frac{d}{du} P\left(Y \geq \frac{1-u}{u}x \,\middle|\, X=x\right)$$

$$= \frac{x}{u^2} f_{Y|X}\left(\frac{1-u}{u}x \,\middle|\, x\right) \overset{\text{(indep)}}{=} \frac{x}{u^2} f_Y\left(\frac{1-u}{u}x\right)$$

$$= \frac{x}{u^2} \cdot \frac{1}{\Gamma(\beta)} \left(\frac{1-u}{u}x\right)^{\beta-1} e^{-\frac{1-u}{u}x}.$$

So

$$f_U(u) = \int_0^\infty \overbrace{f_{U|X}(u|x) f_X(x)}^{= f_{U,X}(u,x)}\,dx = \int_0^\infty \frac{(1-u)^{\beta-1}}{u^{\beta+1}} \frac{x^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} e^{-\frac{x}{u}}\,dx$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u^{\alpha-1}(1-u)^{\beta-1}.$$

$\square$

$\boxed{2}$ 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

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

Ex 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, $Y \sim$ Binomial(n, p).

- ▷ $E(Y) = n E(X_1) = np$
- ▷ $Var(Y) = n Var(X_1) = np(1-p)$,
- ▷ $M_Y(t) = M_{X_1}(t)^n = (1-p+pe^t)^n$. ▯

Ex If $X_1, \cdots, X_k$ are indep. and $X_i \sim$ Binomial(n_i, p) for each $i = 1, \cdots, k$, then $Y = X_1 + \cdots + X_k$ satisfies

$$M_Y(t) = M_{X_1}(t) \cdots M_{X_k}(t) = (1-p+pe^t)^{n_1 + \cdots + n_k}$$ ▯

$$\Rightarrow Y \sim \text{Binomial}(n_1 + \cdots + n_k, p).$$

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

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

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

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

▷ 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:

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

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

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

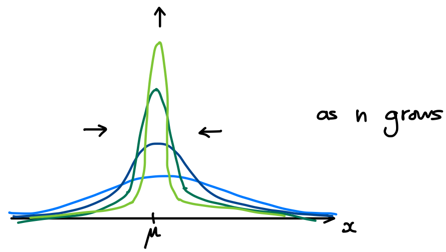
$$\mu_{\overline{X}} = E(\overline{X}) = \frac{1}{n}(\overbrace{\mu + \cdots + \mu}^{n \text{ terms}}) = \mu$$

and

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

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., $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$

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