5. Multivariate Distribution

① Joint Discrete Distribution:

Joint probability: $P[X=x, Y=y] \xrightarrow{\text{def 5.1}} f_{X,Y}(x,y) = P(X=x, Y=y)$

theorem: thm 5.1 A function $f_{X,Y}(x,y)$ is a joint pdf of a 2-dim discrete r.v (X,Y) iff $\begin{cases} f_{X,Y}(x,y) \geq 0 \text{ for all possible } (x,y) \\ \sum_{\text{all } x} \sum_{\text{all } y} f_{X,Y}(x,y) = 1 \end{cases}$

marginal: def 5.2 marginal pdf of X & Y are: $\begin{cases} f_X(x) = \sum_{\text{all } y} f_{X,Y}(x,y) \\ f_Y(y) = \sum_{\text{all } x} f_{X,Y}(x,y) \end{cases}$

CDF def 5.3 CDF of 2-dim r.v (X,Y) is defined as $F_{X,Y}(x,y) = P[X \leq x, Y \leq y]$

$P[X=a, Y=b] = P[X \leq a, Y=b] - P[X \leq a-1, Y=b] = \dots = F_{X,Y}(a,b) - F_{X,Y}(a-1,b) - F_{X,Y}(a,b-1) + F_{X,Y}(a-1,b-1)$

thm 5.2 $F_{X,Y}(x,y)$ is a bivariate CDF for a certain pair (X,Y) iff

i) $\lim_{x \to -\infty} F_{X,Y}(x,y) = F_{X,Y}(-\infty, y) = 0$ for all y

ii) $\lim_{y \to -\infty} F_{X,Y}(x,y) = F_{X,Y}(x, -\infty) = 0$ for all x

iii) $\lim_{x \to \infty, y \to \infty} F_{X,Y}(x,y) = F_{X,Y}(\infty, \infty) = 1$

iv) $F_{X,Y}(b,d) - F_{X,Y}(b,c) - F_{X,Y}(a,d) + F_{X,Y}(a,c) \geq 0$

v) $\lim_{h \downarrow 0} F_{X,Y}(x+h, y) = \lim_{h \downarrow 0} F_{X,Y}(x, y+h) = F_{X,Y}(x,y)$

table 1.

	bivariate	multivariate
	(x,y)	$\vec{x} = (x_1, x_2, \dots, x_k) \rightarrow$ k-dimensional
Joint pdf	$f_{X,Y}(x,y) = P(X=x, Y=y)$	$f_{\vec{X}}(\vec{x}) = P(X_1 = x_1, X_2 = x_2, \dots X_k = x_k)$
theorem	$\begin{cases} f_{X,Y}(x,y) \geq 0 \quad \forall (x,y) \\ \sum_{\text{all } x} \sum_{\text{all } y} f_{X,Y}(x,y) = 1 \end{cases}$	$\begin{cases} f_{\vec{X}}(\vec{x}) \geq 0 \quad \forall \vec{x} \\ \sum_{\text{all } x_1} \dots \sum_{\text{all } x_k} f_{\vec{X}}(\vec{x}) = 1 \end{cases}$
marginal pdf	$f_X(x) = \sum_{\text{all } y} f_{X,Y}(x,y)$; $f_Y(y) = \sum_{\text{all } x} f_{X,Y}(x,y)$	$f_{X_j}(x_j) = \sum_{\text{all } x_1} \dots \sum_{\text{all } x_i \neq x_j} f_{\vec{X}}(\vec{x})$
joint CDF	$F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$	$F_{\vec{X}}(\vec{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \dots X_k \leq x_k)$

table 2

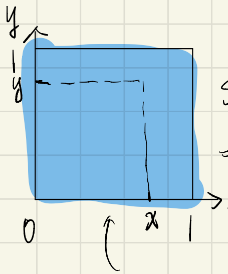
	binomial $X \sim BIN(n,p)$	multinomial $\vec{X} \sim MULT(n, p_1, \dots p_k)$
①	n independent trials	n independent trials
②	each trial has 2 outcomes: S, F	each trial has k+1 possible outcomes: $E_1, \dots E_{k+1}$
③	X = # successes	X_j = "# outcome E_j"; $X_{k+1} = n - X_1 - X_2 \dots - X_k$
④	$p = P(S)$; $q = 1-p = P(F)$	$p_j = P(E_j)$ $j = 1, \dots, k+1$; $p_{k+1} = 1 - \sum_{j=1}^{k} p_j$
⑤	$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, \dots n$	$f_{\vec{X}}(\vec{x}) = \frac{n!}{x_1! x_2! \dots x_{k+1}!} p_1^{x_1} p_2^{x_2} \dots p_{k+1}^{x_{k+1}}$ ①

table 3

	hypergeometric	Multivariate hypergeometric
①	population size N	pop size N
②	2 categories: S, F	k+1 categories, M_i items per categories i $\Rightarrow M_1 + M_2 + \dots M_{k+1} = N$
③	n draws without replacement	n draws with replacement
④	X = num - succ in sample	X_i = number of items of categories i drawn.
⑤	$f_X(x) = \binom{M}{x}\binom{N-M}{n-x} / \binom{N}{n}$, for $x = 0, 1, \dots \min(n, M)$	$f_{\vec{X}}(\vec{x}) = \binom{M_1}{x_1}\binom{M_2}{x_2} \dots \binom{M_{k+1}}{x_{k+1}} / \binom{N}{n}$ for $x_i = 0, 1, \dots M_i$

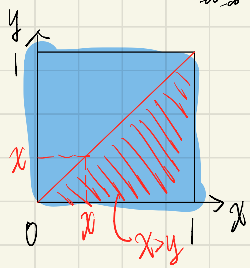
② Joint continuous distributions:

theorem: (th5.4) $f_{X,Y}(x,y)$ is a bivariate pdf for continuous r.v.s (X,Y) iff: $\begin{cases} f_{X,Y}(x,y) \geq 0 \quad \forall (x,y) \in \mathbb{R}^2 \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1 \end{cases}$



Supporter $= \int_0^1\int_0^1 (x^2 + \frac{8}{3}xy)\,dy$

Case 1.



$P(X>Y) = \int_0^1\int_0^x (x^2 + \frac{8}{3}xy)\,dy$

Joint CDF: $F_{X,Y}(x,y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(u,v)\,du\,dv$.

$P(X \leq x, Y \leq y) \rightleftarrows f_{X,Y}(x,y)$ $\quad \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$ $\quad \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(u,v)\,du\,dv$

Marginal pdf: (df5.10) $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$; $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$

③ Expected values of functions of r.v.s:

(X,Y) with joint pdf $f_{X,Y}(x,y)$ & $W = g(X,Y)$ (th5.6)

$E(W) = E[g(X,Y)] = \begin{cases} \sum_{\text{all } x}\sum_{\text{all } y} g(x,y) f_{X,Y}(x,y) & \text{if } X, Y \text{ disc} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy & \text{if } X, Y \text{ conti} \end{cases}$

(th5.7) Let X & Y be r.v.s and let $a,b,c \in \mathbb{R}$ $\quad E[aX+bY+c] = aE[X] + bE[Y] + c$

in particular $\quad E[X+Y] = E[X] + E[Y]$

④ Independence of random variables, covariance & correlation

Independence: r.v X & Y are independent if for every $a<b$ & $c<d$ it holds that:

(df5.13) $P(a \leq X \leq b, c \leq Y \leq d) = P(a \leq X \leq b) \cdot P(c \leq Y \leq d)$

(th5.9) $F_{X,Y}(x,y) = F_X(x)F_Y(y) \ \forall x,y$ / $f_{X,Y}(x,y) = f_X(x)f_Y(y) \ \forall x,y$

(th5.10) X, Y with pdf $f_{X,Y}(x,y)$ are indep. iff: 1) Support of X & Y is Cartesian product of the supports of X & Y respectively, so $\{(x,y) \mid f_{X,Y}(x,y) > 0\} = \{x \mid f_X(x) > 0\} \times \{y \mid f_Y(y) > 0\}$

Condition:

2) $f_{X,Y}(x,y)$ can be written as $g(x)h(y)$ for all values of X & Y belonging to the support of X & Y.

theories: If X, Y are independent:

a) $E(g(X)h(Y)) = E(g(X)) \cdot E(h(Y))$ b) $G_{X+Y}(t) = G_X(t) G_Y(t)$

c) $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$ d) $E(X \cdot Y) = E(X) \cdot E(Y)$ e) $Var(X+Y) = Var(X) + Var(Y)$

Covariance

$$Var(X+Y) = Var(X) + Var(Y) + 2\underbrace{(E(XY) - E(X)E(Y))}_{Cov}$$

(def 5.14) $\sigma_{XY} = Cov(X,Y) = E[(X-E(X))(Y-E(Y))] = E[(X-\mu_X)(Y-\mu_Y)]$

(thm 5.12) $cov(X,Y) = E(XY) - E(X)E(Y)$

$Cov(X,Y) = \sum_{\text{all } x}\sum_{\text{all } y} (x-\mu_X)(y-\mu_Y) f_{X,Y}(x,y)$

theorems:

1) $Cov(aX, bY) = ab\,Cov(X,Y)$

2) $Cov(X+a, Y+b) = Cov(X,Y)$

3) $Cov(X, aX+b) = a\,Var(X)$

4) $Cov(X+U, Y) = Cov(X,Y) + Cov(U,Y)$

(thm 5.14) $Var(aX+bY+c) = a^2 Var(X) + b^2 Var(Y) + 2ab\,Cov(X,Y)$

Correlation coef: The correlation coefficient of X & Y is $\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$ (def 5.15)

(thm 5.15) For any X & Y: $-1 \le \rho_{XY} \le 1$ and $|\rho_{XY}| = 1$ if there exists an exact linear relation ($Y = aX+b$) between X & Y

Only, X, Y are independent $\Rightarrow Cov(X,Y) = \rho_{XY} = 0$

mutual & pairwise independence

(def 5.16) The random variables $(X_1, X_2, X_3, \dots X_n)$ are mutually independent for $a_i < b_i$ we got:

$$P(a_1 \le X_1 \le b_1, a_2 \le X_2 \le b_2, \dots a_k \le X_k \le b_k) = \prod_{j=1}^{k} P(a_j \le X_i \le b_j)$$

mutual → pairwise

(thm 5.16) The r.v $(X_1, X_2, \dots X_k)$ are mutually independent iff:

$$F_{\vec{X}}(\vec{x}) = F_{(X_1,\dots X_k)}(x_1, \dots x_k) = \prod_{i=1}^{k} F_{X_i}(x_i) \quad \forall \vec{x} \in \mathbb{R}^k$$

or $$f_{\vec{X}}(\vec{x}) = f_{X_1,\dots X_k}(x_1, \dots x_k) = \prod_{i=1}^{k} f_{X_i}(x_i) \quad \forall \vec{x} \in \mathbb{R}^k$$

pairwise independence: $f_{X_i, X_j}(x_i, x_j) = f_{X_i}(x_i) f_{X_j}(x_j) \quad \forall x_i \in \mathbb{R}, x_j \in \mathbb{R}$.

(thm 5.17) $M_{X_1+X_2+X_3+\dots+X_k}(t) = \prod_{i=1}^{k} M_{X_i}(t)$

Covariance:

(thm 5.18) $Var[\sum_{i=1}^{k} a_i X_i] = \sum_{i=1}^{k} a_i^2 Var[X_i] + 2\sum_{i<j}\sum a_i a_j Cov(X_i, X_j) = \sum_{i=1}^{k} a_i^2 \sigma_i^2 + 2\sum_{i<j}\sum a_i a_j \sigma_{ij}$

$$Var[\sum_{i=1}^{n} X_i] \overset{\perp}{=} \sum_{i=1}^{n} Var[X_i]$$

Var-Cov matrix:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \dots & \sigma_{1k} \\ \sigma_{21} & & & \\ \sigma_{31} & & & \\ \vdots & & & \\ \sigma_{k1} & \dots & & \sigma_k^2 \end{pmatrix}$$

$$Var[\vec{a}^T \vec{X}] = Var[\vec{a} \cdot \vec{x}] = \vec{a}^T \Sigma \vec{a}$$

⑤ Conditional Distributions:

Conditional pdf: The conditional pdf of y given $X=x$: $f_{Y|X}(y|x) = f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$
(def5.17).

properties:
① $f_{Y|X}(y|x) \geq 0 \ \forall y$
② $\sum_{\text{all } y} f_{Y|X}(y|x) = 1$ or $\int_{-\infty}^{\infty} f_{Y|X}(y|x)dy = 1$ } y is the r.v and
③ $\sum_{\text{all } x} f_{Y|X}(y|x) \neq 1$ or $\int_{-\infty}^{\infty} f_{Y|X}(y|x)dx \neq 1$ } $X=x$ is a parameter.
④ $P(a \leq Y \leq b \mid X=x) = \int_a^b f_{Y|X=x}(y)dy$
⑤ "conditional CDF": $F_{Y|X}(y|x) = P(Y \leq y \mid X=x) = \int_{-\infty}^{y} f_{Y|X}(v|x)dv$.

for indep. (thm5.17) X & Y are indep. r.v, then: $f_{Y|X}(y|x) = f_Y(y) \ \forall y$; $f_{X|Y}(x|y) = f_X(x) \ \forall x$.

$P(A) = P((X,Y) \in A) = \sum\sum_{(x,y)\in A} f_{X,Y}(x,y) = \sum_x P(A \mid X=x) f_X(x)$

Conditional Expected value
$E(Y|X=x) = \begin{cases} \sum_{\text{all } y} y f_{Y|X}(y|x) \\ \int y f_{Y|X}(y|x)dy \end{cases}$
$E(Y|X=x) \neq E[Y]$
$E(g(X,Y)|X=x) = \begin{cases} \sum g(x,y) f_{Y|X}(y|x) \\ \int g(x,y) f_{Y|X}(y|x)dy \end{cases}$

(thm5.21) $E[E[Y|X]] = E[Y]$

Conditional Var (def5.19) The conditional variance of y given $X=x$:
$Var[Y|X=x] = E[(Y-E[Y|X=x])^2 \mid X=x] = E[Y^2|X=x] - E[Y|X=x]^2$

(thm5.24) $Var[Y] = Var[E[Y|X]] + E[Var[Y|X]]$

⑥ Joint moment generating functions:

Joint MGF: (def5.20) Joint MGF of r.v.s $(X_1, X_2, \ldots X_k)$ is $M_{X_1,\ldots X_k}(t_1,t_2,\ldots t_k) = E[e^{t_1X_1+t_2X_2+\ldots+t_kX_k}]$

properties:
① $E[X_i^r X_j^s] = \frac{\partial^{r+s}}{\partial t_i^r \partial t_j^s} M_{\vec{X}}(\vec{t})\Big|_{t_1=t_2=\ldots=t_k=0}$
② Suppose the joint MGF $M_{X,Y}(t_1,t_2)$ for (X,Y) is given. $M_X(t) = M_{X,Y}(t,0)$ and $M_Y(t) = M_{X,Y}(0,t)$
③ Joint MGF of $X_1,\ldots,X_k$ only needs to be defined on an open region $(t_1,\ldots t_k) \in U \subseteq \mathbb{R}^k$ which contains the origin $\vec{0} \in \mathbb{R}^k$
④ "Uniqueness property" $(X_1,\ldots X_k) \sim (Y_1,\ldots Y_k) \Leftrightarrow M_{\vec{X}}(\vec{t}) = M_{\vec{Y}}(\vec{t})$
⑤ (thm5.25) X, Y are independent $\Leftrightarrow M_{X,Y}(t_1,t_2) = M_X(t_1)M_Y(t_2)$

⑦ Bivariate normal distribution:

(Def5.21) $f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]}$

Marginal pdf (thm5.26) If $(X,Y) \sim BVN(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, then $X \sim N(\mu_x,\sigma_x^2)$ and $Y \sim N(\mu_y,\sigma_y^2)$

Conditional pdf

(thm5.27) If $(X,Y) \sim BVN(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$, then $Y|X=x \sim N(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X), \sigma_Y^2(1-\rho^2))$

$X|Y=y \sim N(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y), \sigma_X^2(1-\rho^2))$

Joint MGF.

(thm5.28) If $(X,Y) \sim BVN(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$, then $M_{X,Y}(t_1,t_2) = \exp(\mu_X t_1 + \mu_Y t_2 + \frac{1}{2}[\sigma_X^2 t_1^2 + \sigma_Y^2 t_2^2 + 2\rho\sigma_X\sigma_Y t_1 t_2])$

$\rho = \rho_{X,Y}$.

Independence

$(X,Y) \sim BVN$: X, Y independent $\Leftrightarrow \rho_{X,Y} = 0$

Multivariate normal distribution:

(def5.22) $f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{\frac{k}{2}}|\Sigma|} \exp(-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^T)$

$\left.\begin{array}{l} \Sigma \Rightarrow \text{cov-var matrix} \rightarrow \frac{1}{2}k(k+1) \\ \vec{\mu} \rightarrow (\mu_1, \ldots, \mu_k)^T \rightarrow k \end{array}\right\} \frac{1}{2}k(k+3)$ parameters.