Continuous Random Variables

① Gamma Distribution : pdf, $f(x|\alpha,\beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & \text{for } x>0 \\ 0 & \text{otherwise} \end{cases}$ $\alpha,\beta \in \mathbb{R}^+$.

Properties: 1) $\underbrace{E(X^k) = \frac{\Gamma(\alpha+k)}{\beta^k \Gamma(\alpha)}}_{\text{proof ①.}}$, $Var(X) = \frac{\alpha}{\beta^2}$, $E[X] = \frac{\alpha}{\beta}$.

2) $X_i \sim Gamma(\alpha_i,\beta)$, X_i independent

$\Rightarrow Y = \sum_{i=1}^{k} X_i \sim Gamma\left(\sum_{i=1}^{k} \alpha_i, \beta\right)$

Another parametrization: $f(x|\alpha,\beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$

special case Gamma $(\alpha=\frac{p}{2}, \beta=\frac{1}{2})$

② Chi-square Distribution : χ_p^2 distribution with degrees of freedom p

pdf: $f(x|p) = \frac{1}{\Gamma(\frac{p}{2})\, 2^{\frac{p}{2}}} x^{\frac{p}{2}-1} e^{-\frac{x}{2}}$, $0<x<\infty$.

Property: - $E(X) = \frac{\alpha}{\beta} = p$, $Var(X) = \frac{\alpha}{\beta^2} = 2p$.

- If $Z \sim N(0,1)$, then $Z^2 \sim \chi_1^2$

- If X_i are independent. $X_i \sim \chi^2_{p_i}$, then $\sum_{i=1}^{n} X_i \sim \chi^2_{p_1+\cdots+p_n}$

③ Bivariate Normal Density Function

$X, Y \in \mathbb{R}$. $f(x,y) = \frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right\}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$

where ① μ_x, μ_y are marginal means.

② $\sigma_x, \sigma_y > 0$ are marginal standard deviation.

③ $0 \le |\rho| \le 1$ is the correlation coefficient

- $X \sim N(\mu_x, \sigma_x^2)$, $Y \sim N(\mu_y, \sigma_y^2)$

- if X and Y are bivariate normal, X, Y independent $\boxed{\Leftrightarrow}$ $f(x,y) = f(x)f(y)$

- $Z = aX + bY$, correlation coefficient ρ, Then $Z \sim N(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y)$

$\begin{bmatrix} X \\ Y \end{bmatrix} = N_2\left[\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}\right]$ where $cor(X,Y) = \rho$.

④ Multivariate Normal Distribution.

let $x = (x_1, \cdots, x_p)^T$, let $x \sim N_p(\mu, \Sigma)$.

The multivariate normal density function is $f(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\}$

where - $\mu = (\mu_1, \cdots, \mu_p)^T$ is p-dimensional mean vector

- $\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & & \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{pmatrix}$ is the covariance matrix. $\sigma_{ii} = \sigma_i^2 = Var(X_i)$.

Properties on parameters

- $\mu_j \in \mathbb{R}$ $\forall j$.
- $\sigma_{jj} > 0$ $\forall j$.
- $\sigma_{ij} = \rho_{ij}\sqrt{\sigma_{ii}\sigma_{jj}}$, $\rho_{ij} = Corr(x_i, x_j)$.
- $\sigma_{ij}^2 \le \sigma_{ii}\sigma_{jj}$ $\forall i,j \in \{1, \cdots, p\}$.

$A = \{a_{ij}\}_{n\times p}$ is a non-random matrix, ~~b=~~

$b = (b_1, \cdots b_n)^T$ is non-random vector.

Affine Transformation

Define $y = Ax + b$ $A \neq O_{n\times p}$. Then $y \sim N_n(A\mu + b, A\Sigma A^T)$

*Linear combinations of normal variables are normally distributed

If $a \in \mathbb{R}^p$, $a^T x \sim N(a^T\mu, a^T\Sigma a)$ $a = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ → jth entry

⇓

Thm let $x \sim N(\mu, \Sigma)$. Then $x_j \sim N(\mu_j, \sigma_{jj})$

* Subset of Multivariate Normal Random Variables are multivariate normal

let $z \sim N_p(\mu, \Sigma)$, $z = \begin{pmatrix} x \\ y \end{pmatrix}$, $x_{p_x \times 1}$, $y_{p_y \times 1}$, $\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_x & \Sigma_{yx} \\ \Sigma_{xy} & \Sigma_y \end{pmatrix}$

⇒ Distributions of x and y: $X \sim N(\mu_X, \Sigma_X)$, $Y \sim N(\mu_Y, \Sigma_Y)$

* $Cov(x_i, x_j) = 0$ ⇒ x_i, x_j independent

Expectation Value and Variance

① Law of Total expectation: $E[X] = E[E[X|Y]]$.

Example ①.

$$E[X] = \begin{cases} \sum_y E[X|Y=y]\,p(Y=y) & \text{for discrete } y \\ \int_{-\infty}^{\infty} E[X|Y=y]\,f(y)\,dy & \text{for continuous } y. \end{cases}$$

② Law of Total Variance:

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

Random Vector and Matrix

random vector: $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, expectation $E[x] = \begin{pmatrix} E[x_1] \\ \vdots \\ E[x_n] \end{pmatrix}$

random matrix $Z = \{Z_{ij}\}$ with expectation $E[Z] = \{E[Z_{ij}]\}$.

Properties of Expectation

a, A, B, C non-random vector and matrix, X, Y random matrices.

- $E[a] = a$, $E[A] = A$.

~~If X and Y are random matrice, then E[X+Y] = E[X]~~ Matrices.

- ~~If~~ $E[X+Y] = E[X] + E[Y]$.
- $E[AX] = AE[X]$
- $E[AXB+C] = AE[X]B + C$

Covariance Matrices.

$$Cov(x) = \begin{pmatrix} Var(x_1) & Cov(x_1, x_2) & \cdots & Cov(x_1, x_n) \\ Cov(x_2, x_1) & & & \vdots \\ \vdots & & & \\ Cov(x_n, x_1) & & \cdots & Var(x_n) \end{pmatrix}$$

$$Cov(X) = E[(x - E[x])(x - E[x])^T]$$

* If $x_1, \cdots, x_n$ are independent, then $Cov(x)$ is diagonal

* If $Cov(x)$ is diagonal, this implies that $x_1, \cdots, x_n$ are uncorrelated, Not independent

Properties of Covariance Matrices

x, y: n- dimensional random vector

a : constant non-random vector

A, B: constant non-random matrix.

- $Cov(x) = Cov(x)^T$
- $Cov(a+x) = Cov(x)$
- $Cov(Ax) = A\, Cov(x) A^T$ proof ②
- Covariance matrices are positive definite.
- $Cov(Ax, By) = A\, Cov(x,y) B^T$

Covariance of Two Random Vectors

$$Cov(x,y) = \begin{pmatrix} Cov(x_1, y_1) & \cdots & Cov(x_1, y_n) \\ \vdots & & \vdots \\ Cov(x_m, y_1) & \cdots & Cov(x_m, y_n) \end{pmatrix}_{m \times n}$$

$$Cov(x,y) = E[(x - E(x))(y - E[y])^T].$$

plus: $Cov(x, x) = Cov(x)$.

另: Partitioned Covariance Matrix

let $z = \begin{pmatrix} x \\ y \end{pmatrix}$, Then $Cov(z) = \begin{pmatrix} Cov(x) & Cov(x,y) \\ Cov(y,x) & Cov(y) \end{pmatrix}$

Proof ①: $E[X^k] = \int_0^\infty x^k f(x|\alpha,\beta)\,dx$

$$= \int_0^\infty x^k \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\,dx$$

$$= \int_0^\infty \underbrace{x^{\alpha+k-1} e^{-\beta x}}_{\text{kernel of Gamma}(\alpha+k,\beta).} \frac{\beta^\alpha}{\Gamma(\alpha)}\,dx$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \frac{\Gamma(\alpha+k)}{\beta^{\alpha+k}} \cdot \underbrace{\frac{\beta^{\alpha+k}}{\Gamma(\alpha+k)} x^{\alpha+k-1} e^{-\beta x}}_{f(x|\alpha+k,\beta)}\,dx$$

$$= \frac{\beta^\alpha \Gamma(\alpha+k)}{\Gamma(\alpha)\beta^{\alpha+k}}$$

$$= \frac{\Gamma(\alpha+k)}{\beta^k \Gamma(\alpha)}$$

proof ②:

$$Cov(Ax) = E[(Ax - E[Ax])(Ax - E[Ax])^T]$$

$$= E[(Ax - AE[x])(Ax - AE[x])^T]$$

$$= A\,E[(x - E(x))(x - E(x))^T]\,A^T$$

$$= A\,Cov(x)\,A^T$$

Example ①: Rolling dice

for each roll, you paid the roll value

① If roll $\{4,5,6\}$, roll again

② If roll $\{1,2,3\}$, game stops.

What is the expected pay-off?

$$E[X|Y\in\{1,2,3\}] = 2$$

$$E[X|Y\in\{4,5,6\}] = 5 + E(X)$$

$$E[X] = E[E[X|Y]] = P(Y\in\{1,2,3\})\,E[X|Y\in\{1,2,3\}] + P(Y\in\{4,5,6\})\,E[X|Y\in\{4,5,6\}]$$

$$= 3.5 + \tfrac{1}{2}E(X)$$

$\Rightarrow E(X) = 7$