Joint Distribution Functions

Conditional Distributions

A conditional distribution provides the probability of a random variable, given that it was conditioned on the value of a second random variable.

Discrete Conditional Distributions

Let X and Y be 2 discrete random variables, with a joint distribution function of

$$p_{X,Y}(x, y) = P(X = x, Y = y)$$

The conditional distribution of $X|Y = y$ is defined as

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

Continuous Conditional Distributions

Let X and Y be 2 continuous random variables, with a joint density function of $f_{X,Y}(x,y)$. The conditional distribution of $X|Y=y$ is defined as

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Bivariae Normal Conditional Distribution

$$f(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} * \left[\right.\right.$$

$$\left.\left. * \left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 \right] \right] \right\}$$

$$f(x) = \frac{1}{\sqrt{2\pi \sigma_x^2}} \exp\left\{ -\frac{1}{2} \left(\frac{x-\mu_x}{\sigma_x}\right)^2 \right\}$$

Find $f(y|x) = \dfrac{f(x,y)}{f(x)}$

$$z_x = \left(\frac{x-\mu_x}{\sigma_x}\right) \quad z_y = \left(\frac{y-\mu_y}{\sigma_y}\right) \quad \gamma = \frac{1}{2(1-\rho^2)}$$

$$k^{(xy)} = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \qquad k^{(x)} = \frac{1}{\sqrt{2\pi\sigma_x^2}} \qquad k = \frac{k^{(xy)}}{k^{x}}$$

$$k\,\frac{\exp\{-\gamma(z_x^2 - 2\rho z_x z_y + z_y^2)\}}{\exp\{-\frac{1}{2}z_x^2\}}$$

$$k\exp\{-\gamma(z_x^2 - 2\rho z_x z_y + z_y^2) + \frac{1}{2}z_x^2\}$$

$$k\exp\{-\gamma z_x^2 + 2\gamma\rho z_x z_y - \gamma z_y^2 + \frac{1}{2}z_x^2\}$$

$$k\exp\{-\gamma z_y^2 + 2\gamma\rho z_x z_y + z_x^2(\frac{1}{2} - \gamma)\}$$

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

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

$$\frac{-\rho^2}{2(1-\rho^2)} \rightarrow -\rho^2\gamma$$

$$k\exp\left\{-\gamma z_y^2 + 2\rho\gamma z_x z_y + z_x^2(-\rho^2\gamma)\right\}$$

$$\downarrow$$

$$K\exp\left\{-\gamma\left[z_y^2 - 2z_y z_x\rho + (z_x\rho)^2\right]\right\}$$

$$k \exp\left\{-\gamma (z_y - z_x \rho)^2\right\}$$

$$k \exp\left\{\frac{-(z_y - z_x \rho)^2}{2(1-\rho^2)}\right\}$$

$$k = \frac{1}{\sqrt{2\pi \sigma_y^2 (1-\rho^2)}}$$

$$\frac{1}{\sqrt{2\pi \sigma_y^2 (1-\rho^2)}} \exp\left\{\frac{-(z_y - z_x \rho)^2}{2(1-\rho^2)}\right\}$$

$$Y|X \sim N(\quad , \sigma_y^2 (1-\rho^2))$$

$$\frac{1}{\sqrt{2\pi\, \sigma_Y^2 (1-\rho^2)}} \exp\left\{ -\frac{(z_Y - z_X\rho)^2}{2(1-\rho^2)} \right\}$$

$$-\frac{(z_Y - z_X\rho)^2}{2(1-\rho^2)} \cdot \frac{\sigma_Y^2}{\sigma_Y^2}$$

$$-\frac{(z_Y - z_X\rho)^2\, \sigma_Y^2}{2\sigma_Y^2 (1-\rho^2)}$$

$$-\left(\left[z_Y - z_X \rho \right] \sigma_Y \right)^2$$

$$-\left[z_y \sigma_y - z_x \rho \sigma_y\right]^2$$

$$-\left(\frac{y-\mu_y}{\cancel{\sigma_y}}\cancel{\sigma_y} - \left(\frac{x-\mu_x}{\sigma_x}\right)\rho\sigma_y\right)^2$$

$$-\left(y - \underbrace{\mu_y - \frac{\sigma_y}{\sigma_x}\rho(x-\mu_x)}_{\mu_{y|x}}\right)^2$$

$$\frac{1}{\sqrt{2\pi\sigma_y^2(1-\rho^2)}}\exp\left\{-\frac{\left(y-\mu_y-\frac{\sigma_y}{\sigma_x}\rho(x-\mu_x)\right)^2}{2\sigma_y^2(1-\rho^2)}\right\}$$

$$Y \mid X = x \sim \mathcal{N}\left(\mu_y - \frac{\sigma_y}{\sigma_x}\rho(x - \mu_x), \sigma_x^2(1-\rho^2)\right)$$

Multivariate Normal

The normal distribution function can be extended to from univariate (n=1), bivariate (n=2), to a k-dimensional normal distribution function:

$$X \sim N_k(\mu, \Sigma)$$

Vector vector matrix

Multivariate Normal Parameters

Variances (X)

Var(x_1) = σ_1^2

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix} \quad \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{21} & \cdots & \sigma_{k1} \\ \sigma_{21} & \sigma_2^2 & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_k^2 \end{pmatrix}$$

Covariance (X_1, X_2) = σ_{21} = σ_{21}

Multivariate Normal PDF

covariance (x)

random vector

$$f_X(x) = \det(2\pi\Sigma)^{-1/2} \exp\left\{-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right\}$$

mean vector

Marginal Normal Distribution

$$X = \begin{pmatrix} X_1 \\ \vdots \\ \vdots \\ X_k \end{pmatrix}$$

$$X = \begin{pmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \end{pmatrix} \; \mathcal{X}_1 = \begin{pmatrix} X_1 \\ \vdots \\ X_j \end{pmatrix} \; \mathcal{X}_2 = \begin{pmatrix} X_{j+1} \\ \vdots \\ X_k \end{pmatrix}$$

Marginal Normal Distribution

$\varpi \leftarrow$ varpi

ϖ

$$\mu = \begin{pmatrix} \varpi_1 \\ \varpi_2 \end{pmatrix} \; \varpi_1 = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_j \end{pmatrix} \; \varpi_2 = \begin{pmatrix} \mu_{j+1} \\ \vdots \\ \mu_k \end{pmatrix}$$

Marginal Normal Distribution

$\Sigma_{11} \rightarrow j \times j$ matrix

$\Sigma_{22} \rightarrow (k-j) \times (k-j)$ matrix

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{21} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix}$$

$\Sigma_{12} \rightarrow (k-j) \times j$

$\Sigma_{21} \rightarrow j \times (k-j)$

Marginal Normal Distribution

$$\mathcal{X}_1 \sim N_j(\varpi_1, \Sigma_{11})$$

$$\mathcal{X}_2 \sim N_{k-j}(\varpi_2, \Sigma_{22})$$

Conditional Normal Distribution

$$\mathcal{X}_2|\mathcal{X}_1 \sim N_j(\varpi_2+\varSigma_{21}\varSigma_{11}^{-1}(\mathcal{X}_1-\varpi_1), \varSigma_{22}-\varSigma_{21}\varSigma_{11}^{-1}\varSigma_{12})$$

Independent Random Variables

Random variables are considered independent of each other if the probability of one variable does not affect the probability of another variable.

Discrete Independent Random Variables

Let X and Y be 2 discrete random variables, with a joint density function of $p_{X,Y}(x,y)$. X is independent of Y if and only if

$$p_{X,Y}(x,y) = p_X(x)p_Y(y)$$

Continuous Independent Random Variables

Let X and Y be 2 continuous random variables, with a joint density function of $f_{X,Y}(x, y)$. X is independent of Y if and only if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

Matrix Algebra

$$A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$$

$$\det(A) = a_1 a_2$$

$$A^{-1} = \begin{pmatrix} 1/a_1 & 0 \\ 0 & 1/a_2 \end{pmatrix}$$

Example

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2 \left\{ \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \right\}$$

Show that $X \perp Y$ (**this will be on the upcoming test!**).

independent

Covariance

Let X and Y be 2 random variables with mean $E(X) = \mu_x$ and $E(Y) = \mu_y$, respectively. The covariance of X and Y is defined as

$$Cov(X, Y) = E\{(X - \mu_x)(Y - \mu_y)\}$$

$$Cov(X, Y) = E(XY) - \mu_x \mu_y$$

Covariance

If X and Y are independent random variables, then

$$Cov(X, Y) = 0$$

Correlation

The correlation of X and Y is defined as

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

Expectations

Let $X = (X_1, X_2, \dots, X_n)^{\mathrm{T}}$ be a set of random variables, the expectation of a function $g(X)$ is defined as

$$E\{g(X)\} = \sum_{x_1 \in X_1} \cdots \sum_{x_n \in X_n} g(X) p(x, \theta)$$

or

$$E\{g(X)\} = \int_{x_1 \in X_1} \cdots \int_{x_n \in X_n} g(X) f(x, \theta) dx_n \cdots dx_1$$

Expected Value and Variance of Linear Functions

Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be random variables with $E(X_i) = \mu_i$ and $E(Y_j) = \tau_j$. Furthermore, let $U = \sum_{i=1}^{n} a_i X_i$ and $V = \sum_{j=1}^{m} b_j Y_j$ where $\{a_i\}_{i=1}^{n}$ and $\{b_j\}_{j=1}^{m}$ are constants. We have the following properties:

- $E(U) = \sum_{i=1}^{n} a_i \mu_i$
- $Var(U) = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2\sum\sum_{i<j} a_i a_j Cov(X_i, X_j)$
- $Cov(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{m} Cov(X_i, Y_j)$

Conditional Expectations

Let X_1 and X_2 be two random variables, the conditional expectation of $g(X_1)$, given $X_2 = x_2$, is defined as

$$E\{g(X_1)|X_2 = x_2\} = \sum_{x_1} g(x_1)p(x_1|x_2)$$

or

$$E\{g(X_1)|X_2 = x_2\} = \int_{x_1} g(x_1)f(x_1|x_2)dx_1.$$

Conditional Expectations

Furthermore,

$$E(X_1) = E_{X_2}\{E_{X_1|X_2}(X_1|X_2)\}$$

and

$$Var(X_1) = E_{X_2}\{Var_{X_1|X_2}(X_1|X_2)\} + Var_{X_2}\{E_{X_1|X_2}$$