MathStats

Densities

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)\,dy}$$

$$F_{Y|X}(y|x) = \int_{-\infty}^{y} \frac{f(x,v)}{f_X(x)}\,dv$$

Expectation

$$\mathbb{E}[x^n] = \int_{-\infty}^{\infty} x^n f_X(x)$$

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} \mathbb{E}[Y|X=x] f_X(x)\,dx$$

$$\mathbb{E}[X^n] = \sum_{x: f(x)>0} x^n f(x)$$

Conditional expectation

$$\mathbb{E}[Y|X=x] = \int y f_{y|x}(y|x) dy$$

$$\mathbb{E}[g(Y)|X=x] = \int g(y) f_{y|x}(y|x) dy$$

Law of iterated expectation

$$\mathbb{E}_X[E_{Y|X}\{g(Y)|X\}] = E[g(Y)]$$

If $X \perp\!\!\!\perp Y$, then

$$E[g(Y)|X=x] = E[g(Y)]$$

Not well defined (cauchy)

$$E[X] = E[X_+] - E[X_-] = \infty - \infty$$

Well defined (cauchy)

$$E[|X|] = \infty$$

Basics

$$\mathrm{Cov}[X,Y] = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$\rho(X,Y) = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\cdot\mathrm{Var}[Y]}} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\sigma_X\sigma_Y}}$$

Expectation Algebra

$$\mathbb{E}[x^n] = \int_{-\infty}^{\infty} x^n f_x(x)$$

Variance Algebra

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + 2\,\mathrm{Cov}[X,Y] + \mathrm{Var}[Y]$$
$$\mathrm{Var}[X-Y] = \mathrm{Var}[X] - 2\,\mathrm{Cov}[X,Y] + \mathrm{Var}[Y]$$
$$\mathrm{Var}[XY] = \mathbb{E}[X^2]\cdot\mathbb{E}[Y^2] - (\mathbb{E}[X]\cdot\mathbb{E}[Y])^2$$
$$\mathrm{Var}[X/Y] = \mathrm{Var}[X\cdot(1/Y)] = \mathrm{Var}[(1/Y)\cdot X]$$
$$\mathrm{Var}[X] = \mathrm{Cov}(X,X) = E[X^2] - E[X]^2$$
$$\mathrm{Var}[aX+bY] = a^2\,\mathrm{Var}[X] + b^2 Var[Y] + 2ab\,\mathrm{Cov}[X,Y]$$

Correlation

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{Var(Y)}}$$

Calc

By parts

$$\int u\,dv = uv - \int v\,du$$

Chain rule

$$f(g(x))' = f'(g(x))g'(x)$$

Product rule

$$(u\cdot v)' = u'\cdot v + u\cdot v'$$

$$(\frac{f(x)}{g(x)})' = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

Jacobian

$$\mathbb{J} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

$$\iint_A g(x,y)\,dx\,dy = \iint_B g(x(u,v), y(u,v))|J(u,v)|\,du\,dv$$

Distributions

Distributions arising from the Normal

$$\Gamma(\frac{1}{2}, \frac{k}{2}) \to \chi^2_k \to t_k \overset{D}{=} \frac{\mathcal{N}(O,1)}{\sqrt{\chi^2_k/k}} \to F_{m,k} \overset{D}{=} \frac{\chi^2_m/m}{\chi^2_k/k}$$

$$F_{1,k} \overset{D}{=} t_k^2$$

Gamma

$$X \sim \Gamma(\lambda, p) \iff \lambda X \sim \Gamma(1,p)$$

$$X \sim \Gamma(\lambda,p) \overset{c>0}{\Longrightarrow} cX \sim \Gamma(\frac{\lambda}{c}, p)$$

$$Theorem(\Gamma(\lambda,p) + \Gamma(\lambda,q) \overset{\perp\!\!\!\perp}{=} \Gamma(\lambda, p+q)):$$

$$\begin{cases} X_1 \sim \Gamma(\lambda,p) \\ X_2 \sim \Gamma(\lambda,q) \\ X_1 \perp\!\!\!\perp X_2 \end{cases} \implies \begin{cases} Y_1 = X_1 + X_2 \sim \Gamma(\lambda, p+q) \\ Y_2 = \frac{X_1}{X_1+X_2} \sim \mathrm{Beta}(p,q) \\ Y_1 \perp\!\!\!\perp Y_2 \end{cases}$$

Chi Squared

Definition

if $\{Z_1,\dots,Z_k\} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, then the distribution of

$$V = Z_1^2 + \cdots + Z_k^2$$

is called the χ^2_k distribution with k degrees of freedom PDF of chi^2_k:

$$f_V(v) = \frac{v^{(k-2)/2}e^{-v/2}}{2^{k/2}\Gamma(k/2)} \sim \Gamma(\frac{1}{2}, \frac{k}{2}), v \geq 0$$

Chi squared distribution with k=2 is a gamma/exp with the following params

$$\chi^2_2 = \Gamma(\frac{1}{2}, 1) = \mathrm{Exp}(\frac{1}{2})$$

T Distribution

if

$$Z \sim \mathcal{N}(0,1)$$
$$V \sim \chi^2_k$$
$$Z \perp\!\!\!\perp V$$

then

$$Q = \frac{Z}{\sqrt{\frac{V}{k}}}$$

which is the Student's t distribution with k degrees of freedom
PDF of t_k:

$$f_Q(q) = c_1(1 + \frac{q^2}{k})^{-(k+1)/2}, q \in \mathbb{R}$$

where $c_1 > 0$ is a constant

F Distribution

$$V \sim \chi^2_m$$
$$W \sim \chi^2_k$$
$$V \perp\!\!\!\perp W$$

Then the distribution of

$$S = \frac{V/m}{W/k}$$

is called the Fisher distribution with m (numerator) and k (denominator) degrees of freedom S has the PDF:

$$f_S(s) = c_2 s^{(m-2)/2}(1 + \frac{m}{k}s)^{-(k+m)/2}, s > 0$$

Special case: if $m=1$, then

$$F_{m,k} = F1, k \overset{D}{=} t_k^2$$

Bivariate Normal

if

$$\mathbf{X} \sim \mathcal{N}(\mu, \Sigma)$$

then a bivariate normal vector $\mathbf{X} = (X_1, X_2)^T$ has the pdf

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma-1(\mathbf{x}-\mu)}$$

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

linear transforms

$$C\mathbf{X} + \mathbf{v} \sim \mathcal{N}(C\mu, C\Sigma C^T)$$
$$\mathbb{E}[\mathbf{X}] = C\boldsymbol{\mu}$$

For $\mathbf{X} = (X_1, X_2)^T \sim$ bivariate Normal, if $\rho = 0$, then

$$cov(X_1, X_2) = 0 \iff X_1 \perp\!\!\!\perp X_2$$

Conditional pdf

$$X_2|X_1 = x_1 \sim \mathcal{N}(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), \sigma_2^2(1-\rho^2))$$

And by definition

$$f_{X_2|X_1}(x_2|x_1) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)}$$

Regression line

$$\mathbb{E}[X_2|X_1] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X_1 - \mu_1)$$

Multivariate Normal

$$\mathbf{X} = \begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$$

$$\mathbf{X_2}|\mathbf{X_1} \sim \mathcal{N}(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{X_1} - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

Convolution

If $X \perp\!\!\!\perp Y$, then $Z = X + Y$ has the PDF

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy$$

Stein's Lemma

If:
function $g(z)$ is differentiable
and $\lim_{z\to-\infty} g(z)\phi(z) = 0$ and $\lim_{z\to+\infty} g(z)\phi(z) = 0$
Then for $Z \sim \mathcal{N}(0,1)$:

$$\mathbb{E}[Zg(Z)] = \mathrm{Cov}(Z, g(Z)) = \mathbb{E}[g'(Z)]$$

Applications
For $Z \sim \mathcal{N}(0,1)$:
$$\mathbb{E}[Z^{2k}] = (2k-1)!!$$
if $\lim_{z\to\pm\infty} m(\mu+\sigma z)\phi(z) = 0$, then:
$$Cov(X, m(X)) = \sigma^2 \mathbb{E}[m'(X)]$$

CLT

Suppose $\{X_1, \ldots, X_n\} \overset{i.i.d.}{\sim} X$, with $\mu = \mathbb{E}[X]$ and $\sigma^2 = \text{Var}(X) \in (0, \infty)$.
Define $Z_n = \sqrt{n}\frac{\overline{X}_n - \mu}{\sigma}$. Then
$$Z_n \xrightarrow{D} Z \sim \mathcal{N}(0,1)$$

LLN

Suppose $\{X_1, \ldots, X_n\} \overset{i.i.d.}{\sim} X$, with $\mu = \mathbb{E}[X]$. Then
$$\overline{X}_n \xrightarrow{D} \mu$$