A Brief Review of Probability

Patrick Lam

Outline

Outline

Expectation

Expectation

The expected value of a random variable X is simply the weighted average of all possible values of X.

Expectation

The expected value of a random variable X is simply the weighted average of all possible values of X.

Discrete Case:

Expectation

The expected value of a random variable X is simply the weighted average of all possible values of X.

Discrete Case:

$$E(X) = \sum_i x_i P(X = x_i)$$

where $P(X = x)$ is the probability mass function (PMF).

Expectation

The expected value of a random variable X is simply the weighted average of all possible values of X.

Discrete Case:

$$E(X) = \sum_i x_i P(X = x_i)$$

where $P(X = x)$ is the probability mass function (PMF).

Continuous Case:

Expectation

The expected value of a random variable X is simply the weighted average of all possible values of X.

Discrete Case:

$$E(X) = \sum_i x_i P(X = x_i)$$

where $P(X = x)$ is the probability mass function (PMF).

Continuous Case:

$$E(X) = \int_{-\infty}^{\infty} x p(x) dx$$

where $p(x)$ is the probability density function (PDF).

Expectation of a Function of a Random Variable

Expectation of a Function of a Random Variable

Suppose we want to find $E[g(X)]$, where $g(X)$ is any function of X.

Expectation of a Function of a Random Variable

Suppose we want to find $E[g(X)]$, where $g(X)$ is any function of X. We can simply weight the values of $g(x)$ by the PDF or PMF of X:

Expectation of a Function of a Random Variable

Suppose we want to find $E[g(X)]$, where $g(X)$ is any function of X. We can simply weight the values of $g(x)$ by the PDF or PMF of X:

$$E[g(X)] = \sum_i g(x_i) P(X = x_i)$$

for discrete random variables

Expectation of a Function of a Random Variable

Suppose we want to find $E[g(X)]$, where $g(X)$ is any function of X. We can simply weight the values of $g(x)$ by the PDF or PMF of X:

$$E[g(X)] = \sum_i g(x_i) P(X = x_i)$$

for discrete random variables and

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) p(x) dx$$

for continuous random variables.

Expectation of a Function of a Random Variable

Suppose we want to find $E[g(X)]$, where $g(X)$ is any function of X. We can simply weight the values of $g(x)$ by the PDF or PMF of X:

$$E[g(X)] = \sum_i g(x_i) P(X = x_i)$$

for discrete random variables and

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) p(x) dx$$

for continuous random variables.

This is sometimes known as the *Law of the Unconscious Statistician* (LOTUS).

Variance

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

We can find the variance using LOTUS,

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

We can find the variance using LOTUS, or we can simplify the formula first.

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

We can find the variance using LOTUS, or we can simplify the formula first.

$$\mathrm{Var}(X) \quad = \quad E[(X - E(X))^2]$$

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

We can find the variance using LOTUS, or we can simplify the formula first.

$$\begin{aligned} \mathrm{Var}(X) &= E[(X - E(X))^2] \\ &= E[X^2 - 2XE(X) + (E(X))^2] \end{aligned}$$

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

We can find the variance using LOTUS, or we can simplify the formula first.

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X - E(X))^2] \\
&= E[X^2 - 2XE(X) + (E(X))^2] \\
&= E(X^2) - 2E(X)E[E(X)] + E([E(X)]^2)
\end{aligned}$$

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

We can find the variance using LOTUS, or we can simplify the formula first.

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X - E(X))^2] \\
&= E[X^2 - 2XE(X) + (E(X))^2] \\
&= E(X^2) - 2E(X)E[E(X)] + E([E(X)]^2) \\
&= E(X^2) - 2[E(X)]^2 + [E(X)]^2
\end{aligned}$$

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

We can find the variance using LOTUS, or we can simplify the formula first.

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X - E(X))^2] \\
&= E[X^2 - 2XE(X) + (E(X))^2] \\
&= E(X^2) - 2E(X)E[E(X)] + E([E(X)]^2) \\
&= E(X^2) - 2[E(X)]^2 + [E(X)]^2 \\
&= \mathbf{E(X^2) - [E(X)]^2}
\end{aligned}$$

Variance

The formula for the variance of a random variable is

$$\mathrm{Var}(X) = E[(X - E(X))^2]$$

We can find the variance using LOTUS, or we can simplify the formula first.

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X - E(X))^2] \\
&= E[X^2 - 2XE(X) + (E(X))^2] \\
&= E(X^2) - 2E(X)E[E(X)] + E([E(X)]^2) \\
&= E(X^2) - 2[E(X)]^2 + [E(X)]^2 \\
&= \mathbf{E}(\mathbf{X}^2) - [\mathbf{E}(\mathbf{X})]^2
\end{aligned}$$

We can then find the first part with LOTUS.

Marginal, Conditional, and Joint Densities

Marginal, Conditional, and Joint Densities

$$p(x) = \int p(x, y) dy$$

Marginal, Conditional, and Joint Densities

$$
\begin{aligned}
p(x) &= \int p(x, y) dy \\
p(x, y) &= \int p(x, y, z) dz
\end{aligned}
$$

Marginal, Conditional, and Joint Densities

$$
\begin{aligned}
p(x) &= \int p(x, y)dy \\
p(x, y) &= \int p(x, y, z)dz \\
p(x|y) &= \frac{p(x, y)}{p(y)}
\end{aligned}
$$

Marginal, Conditional, and Joint Densities

$$\begin{aligned}
p(x) &= \int p(x,y)dy \\
p(x,y) &= \int p(x,y,z)dz \\
p(x|y) &= \frac{p(x,y)}{p(y)} \\
p(x|y,z) &= \frac{p(x,y,z)}{p(y,z)}
\end{aligned}$$

Marginal, Conditional, and Joint Densities

$$
\begin{aligned}
p(x) &= \int p(x,y)dy \\
p(x,y) &= \int p(x,y,z)dz \\
p(x|y) &= \frac{p(x,y)}{p(y)} \\
p(x|y,z) &= \frac{p(x,y,z)}{p(y,z)} \\
p(x,y) &= p(x|y)p(y)
\end{aligned}
$$

Marginal, Conditional, and Joint Densities

$$
\begin{aligned}
p(x) &= \int p(x,y)dy \\
p(x,y) &= \int p(x,y,z)dz \\
p(x|y) &= \frac{p(x,y)}{p(y)} \\
p(x|y,z) &= \frac{p(x,y,z)}{p(y,z)} \\
p(x,y) &= p(x|y)p(y) \\
&= p(y|x)p(x)
\end{aligned}
$$

Marginal, Conditional, and Joint Densities

$$
\begin{aligned}
p(x) &= \int p(x, y) dy \\
p(x, y) &= \int p(x, y, z) dz \\
p(x|y) &= \frac{p(x, y)}{p(y)} \\
p(x|y, z) &= \frac{p(x, y, z)}{p(y, z)} \\
p(x, y) &= p(x|y)p(y) \\
&= p(y|x)p(x) \\
p(x, y, z) &= p(x|y, z)p(y|z)p(z)
\end{aligned}
$$

Outline

Outline

Expectation, Variance, and Densities

Important Distributions

Discrete Distributions

Continuous Distributions

The Bernoulli Distribution

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

$y = 0, 1$

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

$y = 0, 1$

probability of success: $\pi \in [0, 1]$

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

$y = 0, 1$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \pi^y (1-\pi)^{(1-y)}$

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

$y = 0, 1$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \pi^y (1 - \pi)^{(1-y)}$

$E(Y) = \pi$

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

$y = 0, 1$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \pi^y (1 - \pi)^{(1-y)}$

$E(Y) = \pi$

$\text{Var}(Y) = \pi(1 - \pi)$

The Bernoulli Distribution

$Y \sim \text{Bernoulli}(\pi)$

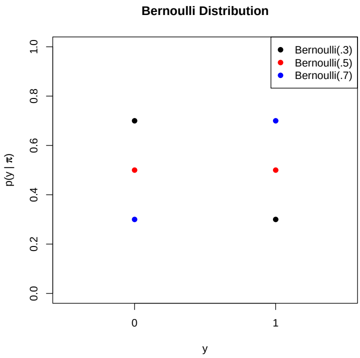
$y = 0, 1$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \pi^y (1 - \pi)^{(1-y)}$

$E(Y) = \pi$

$\text{Var}(Y) = \pi(1 - \pi)$



The Binomial Distribution

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

$y = 0, 1, \ldots, n$

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

$y = 0, 1, \ldots, n$

number of trials: $n \in \{1, 2, \ldots\}$

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

$y = 0, 1, \ldots, n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success: $\pi \in [0, 1]$

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

$y = 0, 1, \ldots, n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \binom{n}{y} \pi^y (1 - \pi)^{(n-y)}$

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

$y = 0, 1, \ldots, n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \binom{n}{y} \pi^y (1 - \pi)^{(n-y)}$

$E(Y) = n\pi$

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

$y = 0, 1, \ldots, n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \binom{n}{y}\pi^y(1-\pi)^{(n-y)}$

$E(Y) = n\pi$

$\text{Var}(Y) = n\pi(1-\pi)$

The Binomial Distribution

$Y \sim \text{Binomial}(n, \pi)$

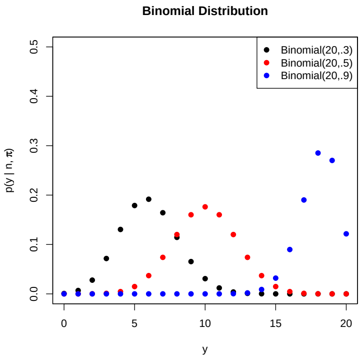
$y = 0, 1, \ldots, n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success: $\pi \in [0, 1]$

$p(y|\pi) = \binom{n}{y} \pi^y (1-\pi)^{(n-y)}$

$E(Y) = n\pi$

$\text{Var}(Y) = n\pi(1-\pi)$



The Multinomial Distribution

The Multinomial Distribution

$$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$$

The Multinomial Distribution

$$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$$

$$y_j = 0, 1, \ldots, n; \;\; \sum_{j=1}^{k} y_j = n$$

The Multinomial Distribution

$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$

$y_j = 0, 1, \ldots, n; \;\; \sum_{j=1}^{k} y_j = n$

number of trials: $n \in \{1, 2, \ldots\}$

The Multinomial Distribution

$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$

$y_j = 0, 1, \ldots, n; \;\; \sum_{j=1}^{k} y_j = n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success for j: $\pi_j \in [0, 1]; \;\; \sum_{j=1}^{k} \pi_j = 1$

The Multinomial Distribution

$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$

$y_j = 0, 1, \ldots, n; \;\; \sum_{j=1}^{k} y_j = n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success for j: $\pi_j \in [0, 1]; \;\; \sum_{j=1}^{k} \pi_j = 1$

$p(\mathbf{y}|n, \boldsymbol{\pi}) = \frac{n!}{y_1! y_2! \ldots y_k!} \pi_1^{y_1} \pi_2^{y_2} \ldots \pi_k^{y_k}$

The Multinomial Distribution

$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$

$y_j = 0, 1, \ldots, n$; $\sum_{j=1}^{k} y_j = n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success for j: $\pi_j \in [0, 1]$; $\sum_{j=1}^{k} \pi_j = 1$

$p(\mathbf{y}|n, \boldsymbol{\pi}) = \frac{n!}{y_1! y_2! \ldots y_k!} \pi_1^{y_1} \pi_2^{y_2} \ldots \pi_k^{y_k}$

$E(Y_j) = n\pi_j$

The Multinomial Distribution

$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$

$y_j = 0, 1, \ldots, n; \;\; \sum_{j=1}^{k} y_j = n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success for j: $\pi_j \in [0, 1]; \;\; \sum_{j=1}^{k} \pi_j = 1$

$p(\mathbf{y}|n, \boldsymbol{\pi}) = \frac{n!}{y_1! y_2! \ldots y_k!} \pi_1^{y_1} \pi_2^{y_2} \ldots \pi_k^{y_k}$

$E(Y_j) = n\pi_j$

$\text{Var}(Y_j) = n\pi_j(1 - \pi_j)$

The Multinomial Distribution

$Y \sim \text{Multinomial}(n, \pi_1, \ldots, \pi_k)$

$y_j = 0, 1, \ldots, n; \;\; \sum_{j=1}^{k} y_j = n$

number of trials: $n \in \{1, 2, \ldots\}$
probability of success for j: $\pi_j \in [0, 1]; \;\; \sum_{j=1}^{k} \pi_j = 1$

$p(\mathbf{y}|n, \boldsymbol{\pi}) = \frac{n!}{y_1! y_2! \ldots y_k!} \pi_1^{y_1} \pi_2^{y_2} \ldots \pi_k^{y_k}$

$E(Y_j) = n\pi_j$

$\text{Var}(Y_j) = n\pi_j(1 - \pi_j)$

$\text{Cov}(Y_i, Y_j) = -n\pi_i\pi_j$

The Poisson Distribution

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

$y = 0, 1, \ldots$

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

$y = 0, 1, \ldots$

expected number of occurrences: $\lambda > 0$

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

$y = 0, 1, \ldots$

expected number of occurrences: $\lambda > 0$

$p(y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

$y = 0, 1, \ldots$

expected number of occurrences: $\lambda > 0$

$p(y|\lambda) = \frac{e^{-\lambda}\lambda^{y}}{y!}$

$E(Y) = \lambda$

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

$y = 0, 1, \ldots$

expected number of occurrences: $\lambda > 0$

$p(y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$

$E(Y) = \lambda$

$\text{Var}(Y) = \lambda$

The Poisson Distribution

$Y \sim \text{Poisson}(\lambda)$

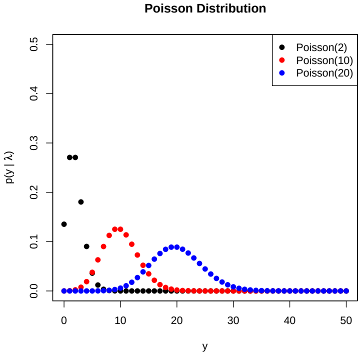
$y = 0, 1, \ldots$

expected number of occurrences: $\lambda > 0$

$p(y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$

$E(Y) = \lambda$

$\text{Var}(Y) = \lambda$



The Geometric Distribution

The Geometric Distribution

How many Bernoulli trials until success?

The Geometric Distribution

How many Bernoulli trials until success?

$Y \sim \text{Geometric}(\pi)$

The Geometric Distribution

How many Bernoulli trials until success?

$Y \sim \text{Geometric}(\pi)$

$y = 1, 2, 3, \ldots$

The Geometric Distribution

How many Bernoulli trials until success?

$Y \sim \text{Geometric}(\pi)$

$y = 1, 2, 3, \ldots$

probability of success: $\pi \in [0, 1]$

The Geometric Distribution

How many Bernoulli trials until success?

$Y \sim \text{Geometric}(\pi)$

$y = 1, 2, 3, \ldots$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = (1 - \pi)^{(y-1)}\pi$

The Geometric Distribution

How many Bernoulli trials until success?

$Y \sim \text{Geometric}(\pi)$

$y = 1, 2, 3, \ldots$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = (1 - \pi)^{(y-1)}\pi$

$E(Y) = \frac{1}{\pi}$

The Geometric Distribution

How many Bernoulli trials until success?

$Y \sim \text{Geometric}(\pi)$

$y = 1, 2, 3, \ldots$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = (1 - \pi)^{(y-1)}\pi$

$E(Y) = \frac{1}{\pi}$

$\text{Var}(Y) = \frac{1-\pi}{\pi^2}$

The Geometric Distribution

How many Bernoulli trials until success?

$Y \sim \text{Geometric}(\pi)$

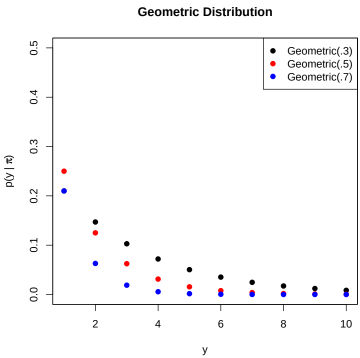
$y = 1, 2, 3, \ldots$

probability of success: $\pi \in [0, 1]$

$p(y|\pi) = (1-\pi)^{(y-1)}\pi$

$E(Y) = \frac{1}{\pi}$

$\text{Var}(Y) = \frac{1-\pi}{\pi^2}$



Outline

The Univariate Normal Distribution

The Univariate Normal Distribution

$Y \sim \text{Normal}(\mu, \sigma^2)$

The Univariate Normal Distribution

$Y \sim \text{Normal}(\mu, \sigma^2)$

$y \in \mathbb{R}$

The Univariate Normal Distribution

$Y \sim \text{Normal}(\mu, \sigma^2)$

$y \in \mathbb{R}$

mean: $\mu \in \mathbb{R}$

The Univariate Normal Distribution

$Y \sim \text{Normal}(\mu, \sigma^2)$

$y \in \mathbb{R}$

mean: $\mu \in \mathbb{R}$
variance: $\sigma^2 > 0$

The Univariate Normal Distribution

$Y \sim \text{Normal}(\mu, \sigma^2)$

$y \in \mathbb{R}$

mean: $\mu \in \mathbb{R}$
variance: $\sigma^2 > 0$

$$p(y|\mu, \sigma^2) = \frac{\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}}$$

The Univariate Normal Distribution

$$Y \sim \text{Normal}(\mu, \sigma^2)$$

$$y \in \mathbb{R}$$

mean: $\mu \in \mathbb{R}$
variance: $\sigma^2 > 0$

$$p(y|\mu, \sigma^2) = \frac{\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}}$$

$$E(Y) = \mu$$

The Univariate Normal Distribution

$Y \sim \text{Normal}(\mu, \sigma^2)$

$y \in \mathbb{R}$

mean: $\mu \in \mathbb{R}$
variance: $\sigma^2 > 0$

$$p(y|\mu, \sigma^2) = \frac{\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}}$$

$E(Y) = \mu$

$\text{Var}(Y) = \sigma^2$

The Univariate Normal Distribution

$$Y \sim \text{Normal}(\mu, \sigma^2)$$

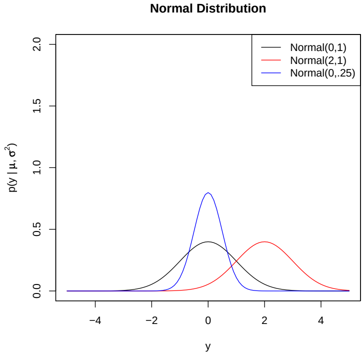
$$y \in \mathbb{R}$$

mean: $\mu \in \mathbb{R}$
variance: $\sigma^2 > 0$

$$p(y|\mu, \sigma^2) = \frac{\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}}$$

$$E(Y) = \mu$$

$$\text{Var}(Y) = \sigma^2$$



The Multivariate Normal Distribution

The Multivariate Normal Distribution

$$Y \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

The Multivariate Normal Distribution

$Y \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\mathbf{y} \in \mathbb{R}^k$

The Multivariate Normal Distribution

$Y \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\mathbf{y} \in \mathbb{R}^k$

mean vector: $\boldsymbol{\mu} \in \mathbb{R}^k$

The Multivariate Normal Distribution

$Y \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\mathbf{y} \in \mathbb{R}^k$

mean vector: $\boldsymbol{\mu} \in \mathbb{R}^k$
variance-covariance matrix: $\boldsymbol{\Sigma}$ positive definite $k \times k$ matrix

The Multivariate Normal Distribution

$Y \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\mathbf{y} \in \mathbb{R}^k$

mean vector: $\boldsymbol{\mu} \in \mathbb{R}^k$
variance-covariance matrix: $\boldsymbol{\Sigma}$ positive definite $k \times k$ matrix

$$p(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\pi}) = (2\pi)^{-k/2}|\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{\mathbf{-1}}(\mathbf{y} - \boldsymbol{\mu})\right)$$

The Multivariate Normal Distribution

$Y \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\mathbf{y} \in \mathbb{R}^k$

mean vector: $\boldsymbol{\mu} \in \mathbb{R}^k$
variance-covariance matrix: $\boldsymbol{\Sigma}$ positive definite $k \times k$ matrix

$p(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\pi}) = (2\pi)^{-k/2}|\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{\mathbf{-1}}(\mathbf{y} - \boldsymbol{\mu})\right)$

$E(Y) = \boldsymbol{\mu}$

The Multivariate Normal Distribution

$Y \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\mathbf{y} \in \mathbb{R}^k$

mean vector: $\boldsymbol{\mu} \in \mathbb{R}^k$
variance-covariance matrix: $\boldsymbol{\Sigma}$ positive definite $k \times k$ matrix

$$p(\mathbf{y}|\boldsymbol{\mu}, \boldsymbol{\pi}) = (2\pi)^{-k/2}|\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{\mathbf{-1}}(\mathbf{y} - \boldsymbol{\mu})\right)$$

$E(Y) = \boldsymbol{\mu}$

$\mathrm{Var}(Y) = \boldsymbol{\Sigma}$

The Uniform Distribution

The Uniform Distribution

$Y \sim \text{Uniform}(\alpha, \beta)$

The Uniform Distribution

$Y \sim \text{Uniform}(\alpha, \beta)$

$y \in [\alpha, \beta]$

The Uniform Distribution

$Y \sim \text{Uniform}(\alpha, \beta)$

$y \in [\alpha, \beta]$

Interval: $[\alpha, \beta]$; $\beta > \alpha$

The Uniform Distribution

$Y \sim \text{Uniform}(\alpha, \beta)$

$y \in [\alpha, \beta]$

Interval: $[\alpha, \beta]$; $\beta > \alpha$

$p(y|\alpha, \beta) = \frac{1}{\beta - \alpha}$

The Uniform Distribution

$Y \sim \text{Uniform}(\alpha, \beta)$

$y \in [\alpha, \beta]$

Interval: $[\alpha, \beta]$; $\beta > \alpha$

$p(y|\alpha, \beta) = \frac{1}{\beta - \alpha}$

$E(Y) = \frac{\alpha+\beta}{2}$

The Uniform Distribution

$Y \sim \text{Uniform}(\alpha, \beta)$

$y \in [\alpha, \beta]$

Interval: $[\alpha, \beta]$; $\beta > \alpha$

$p(y|\alpha, \beta) = \frac{1}{\beta - \alpha}$

$E(Y) = \frac{\alpha + \beta}{2}$

$\text{Var}(Y) = \frac{(\beta - \alpha)^2}{12}$

The Beta Distribution

The Beta Distribution

$$Y \sim \text{Beta}(\alpha, \beta)$$

The Beta Distribution

$Y \sim \text{Beta}(\alpha, \beta)$

$y \in [0, 1]$

The Beta Distribution

$Y \sim \text{Beta}(\alpha, \beta)$

$y \in [0, 1]$

shape parameters:
$\alpha > 0; \;\; \beta > 0$

The Beta Distribution

$Y \sim \text{Beta}(\alpha, \beta)$

$y \in [0, 1]$

shape parameters:
$\alpha > 0; \;\; \beta > 0$

$p(y|\alpha, \beta) =$
$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{(\alpha-1)}(1-y)^{(\beta-1)}$

The Beta Distribution

$Y \sim \text{Beta}(\alpha, \beta)$

$y \in [0, 1]$

shape parameters:
$\alpha > 0; \;\; \beta > 0$

$p(y|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{(\alpha-1)}(1-y)^{(\beta-1)}$

$E(Y) = \frac{\alpha}{\alpha+\beta}$

The Beta Distribution

$Y \sim \text{Beta}(\alpha, \beta)$

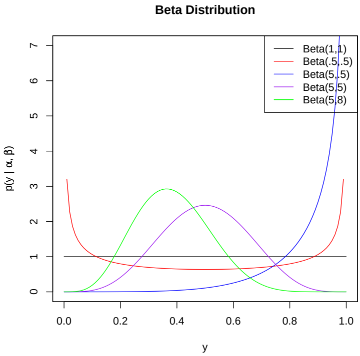
$y \in [0, 1]$

shape parameters:
$\alpha > 0; \quad \beta > 0$

$p(y|\alpha, \beta) =$
$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{(\alpha-1)}(1-y)^{(\beta-1)}$

$E(Y) = \frac{\alpha}{\alpha+\beta}$

$\text{Var}(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2)\alpha+\beta+1)}$



The Gamma Distribution

The Gamma Distribution

$Y \sim \text{Gamma}(\alpha, \beta)$

The Gamma Distribution

$Y \sim \text{Gamma}(\alpha, \beta)$

$y > 0$

The Gamma Distribution

$Y \sim \text{Gamma}(\alpha, \beta)$

$y > 0$

shape parameter: $\alpha > 0$

The Gamma Distribution

$Y \sim \text{Gamma}(\alpha, \beta)$

$y > 0$

shape parameter: $\alpha > 0$
inverse scale parameter: $\beta > 0$

The Gamma Distribution

$Y \sim \text{Gamma}(\alpha, \beta)$

$y > 0$

shape parameter: $\alpha > 0$
inverse scale parameter: $\beta > 0$

$p(y|\alpha, \beta) =$
$\frac{\beta^\alpha}{\Gamma(\alpha)} y^{(\alpha-1)} \exp(-\beta y)$

The Gamma Distribution

$Y \sim \text{Gamma}(\alpha, \beta)$

$y > 0$

shape parameter: $\alpha > 0$
inverse scale parameter: $\beta > 0$

$p(y|\alpha, \beta) =$
$\frac{\beta^\alpha}{\Gamma(\alpha)} y^{(\alpha-1)} \exp\left(-\beta y\right)$

$E(Y) = \frac{\alpha}{\beta}$

The Gamma Distribution

$Y \sim \text{Gamma}(\alpha, \beta)$

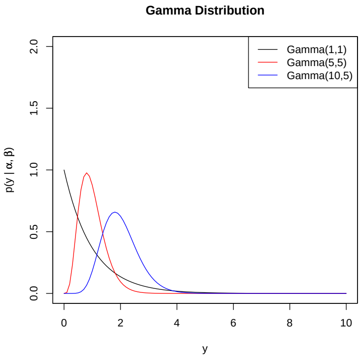
$y > 0$

shape parameter: $\alpha > 0$
inverse scale parameter: $\beta > 0$

$p(y|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{(\alpha-1)} \exp(-\beta y)$

$E(Y) = \frac{\alpha}{\beta}$

$\text{Var}(Y) = \frac{\alpha}{\beta^2}$



The Inverse Gamma Distribution

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution:

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{Invgamma}(\alpha, \beta)$.

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{Invgamma}(\alpha, \beta)$.

$Y \sim \text{Invgamma}(\alpha, \beta)$

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{Invgamma}(\alpha, \beta)$.

$Y \sim \text{Invgamma}(\alpha, \beta)$

$y > 0$

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim$ Gamma(α, β), then $\frac{1}{X} \sim$ Invgamma(α, β).

$Y \sim$ Invgamma(α, β)

$y > 0$

shape parameter: $\alpha > 0$

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{Invgamma}(\alpha, \beta)$.

$Y \sim \text{Invgamma}(\alpha, \beta)$

$y > 0$

shape parameter: $\alpha > 0$
scale parameter: $\beta > 0$

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{Invgamma}(\alpha, \beta)$.

$Y \sim \text{Invgamma}(\alpha, \beta)$

$y > 0$

shape parameter: $\alpha > 0$
scale parameter: $\beta > 0$

$p(y|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} e^{-\frac{\beta}{y}}$

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim$ Gamma(α, β), then $\frac{1}{X} \sim$ Invgamma(α, β).

$Y \sim$ Invgamma(α, β)

$y > 0$

shape parameter: $\alpha > 0$
scale parameter: $\beta > 0$

$p(y|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} e^{-\frac{\beta}{y}}$

$E(Y) = \frac{\beta}{\alpha - 1}$ for $\alpha > 1$

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim$ Gamma(α, β), then $\frac{1}{X} \sim$ Invgamma(α, β).

$Y \sim$ Invgamma(α, β)

$y > 0$

shape parameter: $\alpha > 0$
scale parameter: $\beta > 0$

$p(y|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} e^{-\frac{\beta}{y}}$

$E(Y) = \frac{\beta}{\alpha - 1}$ for $\alpha > 1$

$\text{Var}(Y) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$

The Inverse Gamma Distribution

Distribution of the Inverse of a Gamma Distribution: If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{Invgamma}(\alpha, \beta)$.

$Y \sim \text{Invgamma}(\alpha, \beta)$

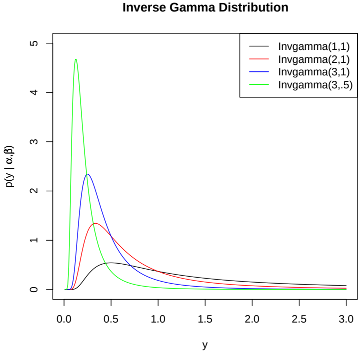
$y > 0$

shape parameter: $\alpha > 0$
scale parameter: $\beta > 0$

$$p(y|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} e^{-\frac{\beta}{y}}$$

$E(Y) = \frac{\beta}{\alpha - 1}$ for $\alpha > 1$

$\text{Var}(Y) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$



The Dirichlet Distribution

The Dirichlet Distribution

$$Y \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_k)$$

The Dirichlet Distribution

$Y \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_k)$

$y_j \in [0, 1]; \;\; \sum_{j=1}^{k} y_j = 1$

The Dirichlet Distribution

$Y \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_k)$

$y_j \in [0, 1]$; $\sum_{j=1}^{k} y_j = 1$

α parameters: $\alpha_j > 0$; $\sum_{j=1}^{k} \alpha_j \equiv \alpha_0$

The Dirichlet Distribution

$Y \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_k)$

$y_j \in [0, 1]; \;\; \sum_{j=1}^{k} y_j = 1$

α parameters: $\alpha_j > 0; \;\; \sum_{j=1}^{k} \alpha_j \equiv \alpha_0$

$p(\mathbf{y}|\boldsymbol{\alpha}) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_k)}{\Gamma(\alpha_1) \ldots \Gamma(\alpha_k)} y_1^{\alpha_1 - 1} \ldots y_k^{\alpha_k - 1}$

The Dirichlet Distribution

$Y \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_k)$

$y_j \in [0, 1]$; $\sum_{j=1}^{k} y_j = 1$

α parameters: $\alpha_j > 0$; $\sum_{j=1}^{k} \alpha_j \equiv \alpha_0$

$p(\mathbf{y}|\boldsymbol{\alpha}) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_k)}{\Gamma(\alpha_1) \ldots \Gamma(\alpha_k)} y_1^{\alpha_1 - 1} \ldots y_k^{\alpha_k - 1}$

$E(Y_j) = \frac{\alpha_j}{\alpha_0}$

The Dirichlet Distribution

$Y \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_k)$

$y_j \in [0, 1]; \;\; \sum_{j=1}^{k} y_j = 1$

α parameters: $\alpha_j > 0; \;\; \sum_{j=1}^{k} \alpha_j \equiv \alpha_0$

$p(\mathbf{y}|\boldsymbol{\alpha}) = \frac{\Gamma(\alpha_1+\cdots+\alpha_k)}{\Gamma(\alpha_1)\ldots\Gamma(\alpha_k)} y_1^{\alpha_1-1} \ldots y_k^{\alpha_k-1}$

$E(Y_j) = \frac{\alpha_j}{\alpha_0}$

$\text{Var}(Y_j) = \frac{\alpha_j(\alpha_0-\alpha_j)}{\alpha_0^2(\alpha_0+1)}$

The Dirichlet Distribution

$Y \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_k)$

$y_j \in [0,1]; \;\; \sum_{j=1}^{k} y_j = 1$

α parameters: $\alpha_j > 0; \;\; \sum_{j=1}^{k} \alpha_j \equiv \alpha_0$

$p(\mathbf{y}|\boldsymbol{\alpha}) = \frac{\Gamma(\alpha_1+\cdots+\alpha_k)}{\Gamma(\alpha_1)\ldots\Gamma(\alpha_k)} y_1^{\alpha_1-1} \ldots y_k^{\alpha_k-1}$

$E(Y_j) = \frac{\alpha_j}{\alpha_0}$

$\text{Var}(Y_j) = \frac{\alpha_j(\alpha_0-\alpha_j)}{\alpha_0^2(\alpha_0+1)}$

$\text{Cov}(Y_i, Y_j) = -\frac{\alpha_i\alpha_j}{\alpha_0^2(\alpha_0+1)}$