Chapter 5 sections

Discrete univariate distributions:

- 5.2 Bernoulli and Binomial distributions
- Just skim 5.3 Hypergeometric distributions
- 5.4 Poisson distributions
- Just skim 5.5 Negative Binomial distributions

Continuous univariate distributions:

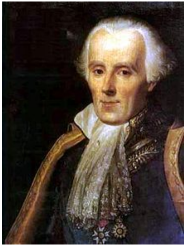
- 5.6 Normal distributions
- 5.7 Gamma distributions
- Just skim 5.8 Beta distributions

Multivariate distributions

- Just skim 5.9 Multinomial distributions
- 5.10 Bivariate normal distributions

Why Normal?

Gauss



Laplace

- Works well in practice. Many physical experiments have distributions that are approximately normal
- Central Limit Theorem: Sum of many i.i.d. random variables are approximately normally distributed
- Mathematically convenient – especially the multivariate normal distribution.
  - Can explicitly obtain the distribution of many functions of a normally distributed random variable have.
  - Marginal and conditional distributions of a multivariate normal are also normal (multivariate or univariate).
- Developed by Gauss and then Laplace in the early 1800s
- Also known at the *Gaussian distributions*

Normal distributions

Def: Normal distributions – $N(\mu, \sigma^2)$

A continuous r.v. X has the *normal distribution with mean* μ *and variance* σ^2 if it has the pdf

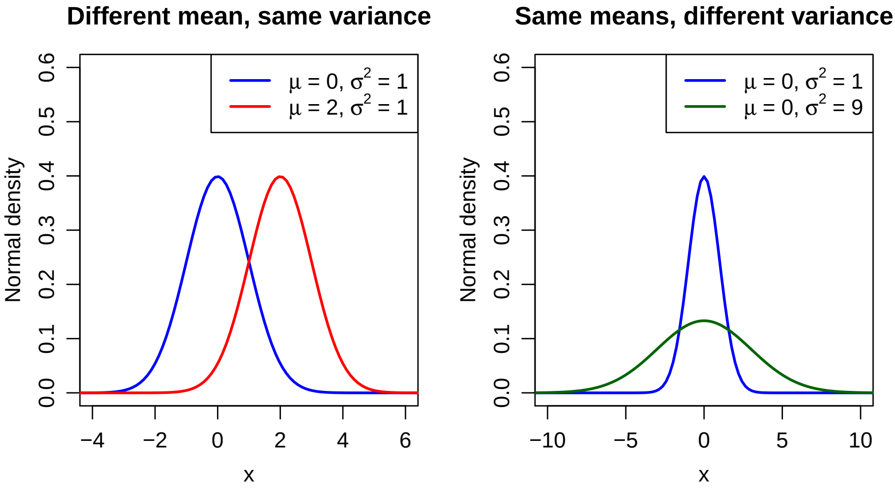
$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty$$

Parameter space: $\mu \in \mathbb{R}$ and $\sigma^2 > 0$

Show:

- $\psi(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$
- $E(X) = \mu$
- $\mathrm{Var}(X) = \sigma^2$

The Bell curve



Standard normal

Standard normal distribution: $N(0, 1)$

The normal distribution with $\mu = 1$ and $\sigma^2 = 1$ is called the *standard normal distribution* and the pdf and cdf are denoted as $\phi(x)$ and $\Phi(x)$

- The cdf for a normal distribution cannot be expressed in closed form and is evaluated using numerical approximations.
- $\Phi(x)$ is tabulated in the back of the book. Many calculators and programs such as R, Matlab, Excel etc. can calculate $\Phi(x)$.
- $\Phi(-x) = 1 - \Phi(x)$
- $\Phi^{-1}(p) = -\Phi^{-1}(1 - p)$

Properties of the normal distributions

Theorem 5.6.4: Linear transformation of a normal is still normal

If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$ where a and b are constants and $a \neq 0$ then
$$Y \sim N(a\mu + b, a^2\sigma^2)$$

- Let F be the cdf of X, where $X \sim N(\mu, \sigma^2)$. Then

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

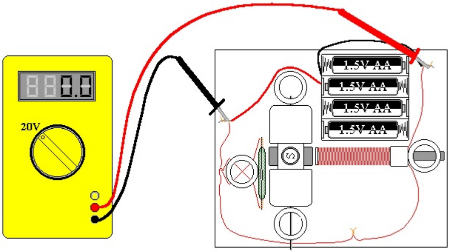
and

$$F^{-1}(p) = \mu + \sigma\Phi^{-1}(p)$$

Example: Measured Voltage

Suppose the measured voltage, X, in a certain electric circuit has the normal distribution with mean 120 and standard deviation 2

1. What is the probability that the measured voltage is between 118 and 122?
2. Below what value will 95% of the measurements be?



Properties of the normal distributions

Theorem 5.6.7: Linear combination of ind. normals is a normal

Let $X_1, \dots, X_k$ be independent r.v. and $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, \dots, k$. Then

$$X_1 + \cdots + X_k \sim N\left(\mu_1 + \cdots + \mu_k,\ \sigma_1^2 + \cdots + \sigma_k^2\right)$$

Also, if $a_1, \dots, a_k$ and b are constants where at least one a_i is not zero:

$$a_1 X_1 + \cdots + a_k X_k + b \sim N\left(b + \sum_{i=1}^{k} \mu_i,\ \sum_{i=1}^{k} a_i^2 \sigma_i^2\right)$$

In particular:

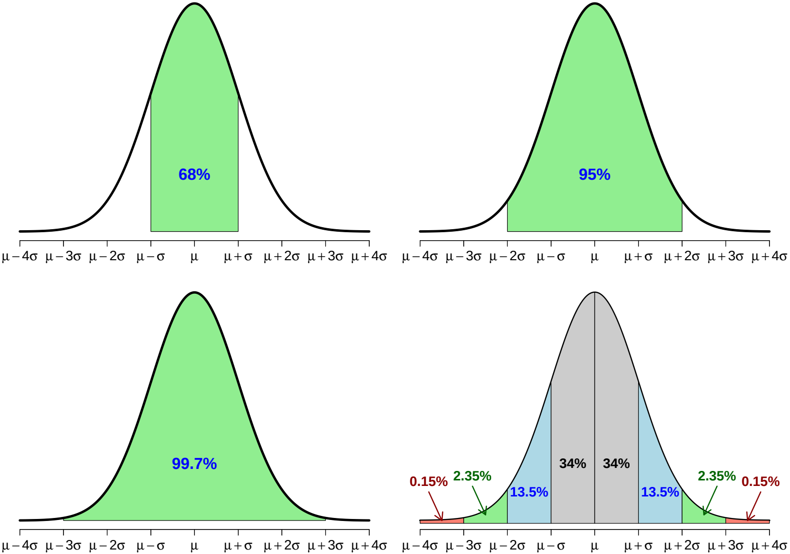
- The *sample mean*: $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$
- If $X_1, \dots, X_n$ are a random sample from a $N(\mu, \sigma^2)$, what is the distribution of the sample mean?

Example: Measured voltage – continued

Suppose the measured voltage, X, in a certain electric circuit has the normal distribution with mean 120 and standard deviation 2.

- If three independent measurements of the voltage are made, what is the probability that the sample mean $\overline{X}_3$ will lie between 118 and 120?
- Find x that satisfies $P(|\overline{X}_3 - 120| \leq x) = 0.95$

Area under the curve

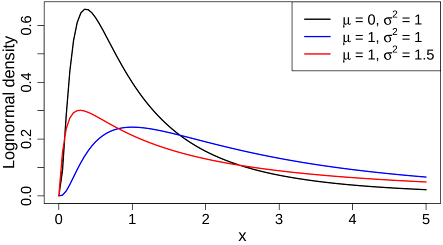


Lognormal distributions

Def: Lognormal distributions

If $\log(X) \sim N(\mu, \sigma^2)$ then we say that X has the *Lognormal distribution with parameters μ and σ^2*.

- The support of the lognormal distribution is $(0, \infty)$.
- Often used to model time before failure.



Example:

- Let X and Y be independent random variables such that $\log(X) \sim N(1.6, 4.5)$ and $\log(Y) \sim N(3, 6)$. What is the distribution of the product XY?

Bivariate normal distributions

Def: Bivariate normal

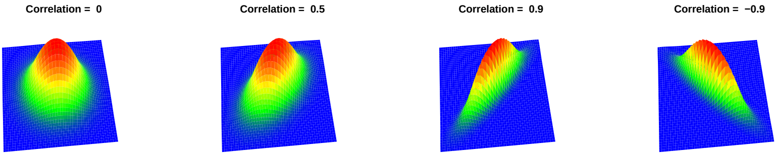
Two continuous r.v. X_1 and X_2 have the *bivariate normal distribution with means μ_1 and μ_2, variances σ_1^2 and σ_2^2 and correlation ρ* if they have the joint pdf

$$f(x_1, x_2) = \frac{1}{2\pi(1-\rho)^{1/2}\sigma_1\sigma_2} \times \exp\left(-\frac{1}{2}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right) \quad (1)$$

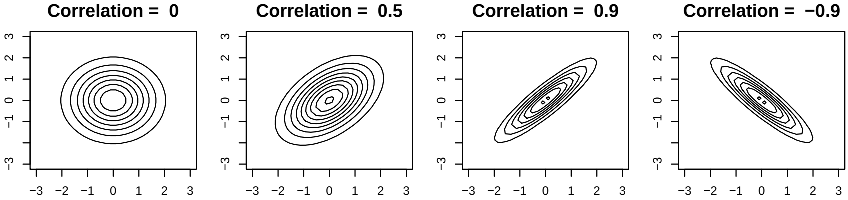
Parameter space: $\mu_i \in \mathbb{R}$, $\sigma_i^2 > 0$ for $i = 1, 2$ and $-1 \leq \rho \leq 1$

Bivariate normal pdf

Bivariate normal pdf with different ρ:



Contours:



Bivariate normal as linear combination

Theorem 5.10.1: Bivariate normal from two ind. standard normals

Let $Z_1 \sim N(0,1)$ and $Z_2 \sim N(0,1)$ be independent.
Let $\mu_i \in \mathbb{R}$, $\sigma_i^2 > 0$ for $i = 1,2$ and $-1 \leq \rho \leq 1$ and let

$$\begin{aligned} X_1 &= \sigma_1 Z_1 + \mu_1 \\ X_2 &= \sigma_2(\rho Z_1 + \sqrt{1-\rho^2} Z_2) + \mu_2 \end{aligned} \qquad (2)$$

Then the joint distribution of X_1 and X_2 is bivariate normal with parameters μ_1, μ_2, σ_1^2, σ_1^2 and ρ

Theorem 5.10.2 (part 1) – the other way

Let X_1 and X_2 have the pdf in (1). Then there exist independent standard normal r.v. Z_1 and Z_2 so that (2) holds.

Properties of a bivariate normal

Theorem 5.10.2 (part 2)

Let X_1 and X_2 have the pdf in (1). Then the marginal distributions are

$$X_1 \sim N(\mu_1, \sigma_1^2) \quad \text{and} \quad X_2 \sim N(\mu_2, \sigma_2^2)$$

And the correlation between X_1 and X_2 is ρ

Theorem 5.10.4: The conditional is normal

Let X_1 and X_2 have the pdf in (1). Then the conditional distribution of X_2 given that $X_1 = x_1$ is (univariate) normal with

$$E(X_2|X_1 = x_1) = \mu_2 + \rho\sigma_2 \frac{(x_1 - \mu_1)}{\sigma_1} \quad \text{and}$$
$$\mathrm{Var}(X_2|X_1 = x_1) = (1 - \rho^2)\sigma_2^2$$

Properties of a bivariate normal

Theorem 5.10.3: Uncorrelated $\Rightarrow$ Independent

Let X_1 and X_2 have the bivariate normal distribution. Then X_1 and X_2 are independent if and only if they are uncorrelated.

- Only holds for the multivariate normal distribution
- One of the very convenient properties of the normal distribution

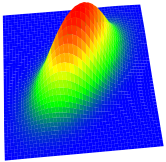
Theorem 5.10.5: Linear combinations are normal

Let X_1 and X_2 have the pdf in (1) and let a_1, a_2 and b be constants. Then $Y = a_1X_1 + a_2X_2 + b$ is normally distributed with

$$E(Y) = a_1\mu_1 + a_2\mu_2 + b \quad \text{and}$$
$$\mathrm{Var}(Y) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2\rho\sigma_1\sigma_2$$

- This extends what we already had for independent normals

Example

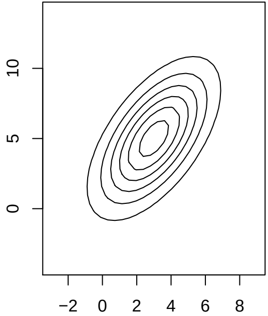


Let X_1 and X_2 have the bivariate normal distribution with means $\mu_1 = 3$, $\mu_2 = 5$, variances $\sigma_1^2 = 4$, $\sigma_2^2 = 9$ and correlation $\rho = 0.6$.

a) Find the distribution of $X_2 - 2X_1$

b) What is expected value of X_2, given that we observed $X_1 = 2$?

c) What is the probability that $X_1 > X_2$?



Multivariate normal – Matrix notation

The pdf of an n-dimensional normal distribution, $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\intercal \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$

where

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} & \cdots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_2^2 & \sigma_{2,3} & \cdots & \sigma_{2,n} \\ \sigma_{3,1} & \sigma_{3,2} & \sigma_3^2 & \cdots & \sigma_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \sigma_{n,3} & \cdots & \sigma_n^2 \end{pmatrix}$$

$\boldsymbol{\mu}$ is the mean vector and Σ is called the *variance-covariance* matrix.

Multivariate normal – Matrix notation

Same things hold for multivariate normal distribution as the bivariate.
Let $\mathbf{X} \sim N(\boldsymbol{\mu}, \Sigma)$

- Linear combinations of **X** are normal
- $A\mathbf{X} + \mathbf{b}$ is (multivariate) normal for fixed matrix A and vector **b**
- The marginal distribution of X_i is normal with mean μ_i and variance σ_i^2
- The off-diagonal elements of Σ are the covariances between individual elements of **X**, i.e. $\mathrm{Cov}(X_i, X_j) = \sigma_{i,j}$.
- The joint marginal distributions are also normal where the mean and covariance matrix are found by picking the corresponding elements from $\boldsymbol{\mu}$ and rows and columns from Σ.
- The conditional distributions are also normal (multivariate or univariate)

Gamma distributions

- The *Gamma function:* $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$
- $\Gamma(1) = 1$ and $\Gamma(0.5) = \sqrt{\pi}$
- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha)$ if $\alpha > 1$

Def: Gamma distributions – $\text{Gamma}(\alpha, \beta)$

A continuous r.v. X has the *gamma distribution with parameters α and β* if it has the pdf

$$f(x|\alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

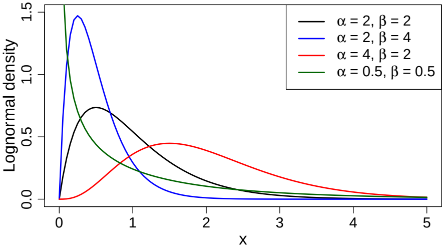
Parameter space: $\alpha > 0$ and $\beta > 0$

- $\text{Gamma}(1, \beta)$ is the same as the *exponential distribution with parameter β*, $\text{Expo}(\beta)$

Properties of the gamma distributions

- $\psi(t) = \left(\frac{\beta}{\beta+t}\right)^{\alpha}$
- $E(X) = \frac{\alpha}{\beta}$ and $E(X) = \frac{\alpha}{\beta^2}$
- If $X_1, \ldots, X_k$ are independent $\Gamma(\alpha_i, \beta)$ r.v. then

$$X_1 + \cdots + X_k \sim Gamma\left(\sum_{i=1}^{k} \alpha_i, \beta\right)$$



Properties of the gamma distributions

Theorem 5.7.9: Exponential distribution is memoryless

Let $X \sim \mathrm{Expo}(\beta)$ and let $t > 0$. Then for any $h > 0$

$$P(X \geq t + h | X \geq t) = P(X \geq h)$$

Theorem 5.7.12: Times between arrivals in a Poisson process

Let Z_k be the time until the k^{th} arrival in a Poisson process with rate β.
Let $Y_1 = Z_1$ and $Y_k = Z_k - Z_{k-1}$ for $k \geq 2$.
Then $Y_1, Y_2, Y_3, \dots$ are i.i.d. with the exponential distribution with parameter β.

Beta distributions

Def: Beta distributions – $\mathrm{Beta}(\alpha, \beta)$

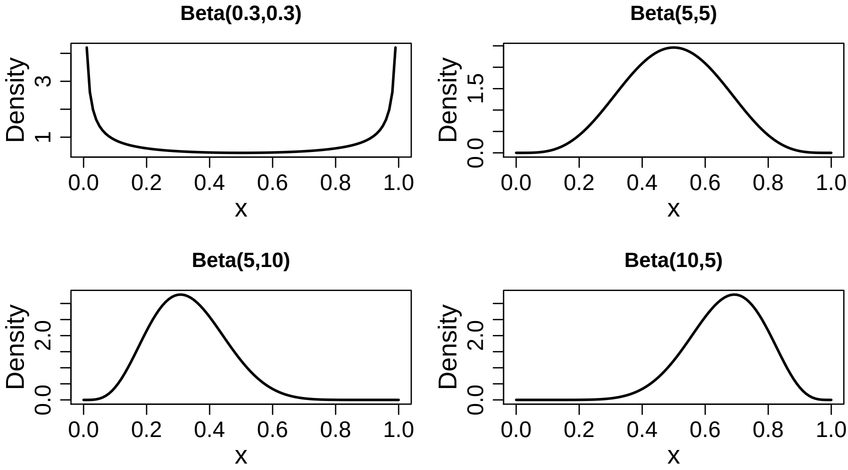
A continuous r.v. X has the *beta distribution with parameters α and β* if it has the pdf

$$f(x|\alpha,\beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Parameter space: $\alpha > 0$ and $\beta > 0$

- $\mathrm{Beta}(1,1) = \mathrm{Uniform}(0,1)$
- Used to model a random variable that takes values between 0 and 1.
- The Beta distributions are often used as *prior distributions* for probability parameters, e.g. the p in the Binomial distribution.

Beta distributions



END OF CHAPTER 5