Chapter 1. Special probability distributions

Reading Assignment

Sections 5.10, 7.2, 8.8, 10.8, 10.9, and Appendix ONE.
Multivariate Normal Distribution

Random vector \( X = (X_1, X_2, \ldots, X_d)' \) has distribution \( N(\mu, \Sigma) \) if the joint density takes form

\[
f(x) = (2\pi)^{-\frac{d}{2}}|\Sigma|^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}.
\]

Here

1. Two parameters: \( \mu = (\mu_1, \ldots, \mu_d)' \) and \( \Sigma = [\sigma_{ij}]_{d \times d} \) is a symmetric positive definite matrix.

2. \( E[X] = \mu, \quad \text{Var}[X] = \Sigma. \)

Or,

\[
E[X_i] = \mu_i, \quad \text{Var}[X_i] = \sigma_{ii}, \quad \text{Cov}(X_i, X_j) = \sigma_{ij} = \sigma_{ji}.
\]
Properties of Multivariate Normal Distribution

1. Suppose $X_1, X_2, \ldots, X_d$ are independent normal random variables with $E[X_i] = \mu_i$ and $\text{Var}[X_i] = \sigma_i^2$.

Then the random vector $X = (X_1, X_2, \ldots, X_d)'$ has (multivariate) normal distribution $N(\mu, \Sigma)$ with

$$
\mu = (\mu_1, \ldots, \mu_d)'
$$

and

$$
\Sigma = [\sigma_{ij}], \quad \sigma_{ii} = \sigma_i^2, \quad \sigma_{ij} = 0 \quad (\text{if } i \neq j)
$$
2. Suppose \( X = (X_1, X_2, \ldots, X_d)' \) has distribution \( N(\mu, \Sigma) \).
Then any marginal distribution of \( X \) is still normal.

For example,
(a) \( X_i \) has distribution \( N(\mu_i, \sigma_{ii}) \).
(b) \( (X_1, X_2) \) has normal distribution \( N(\bar{\mu}, \bar{\Sigma}) \) with

\[
\bar{\mu} = (\mu_1, \mu_2)', \quad \bar{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}
\]
3. Any linear transform of a normal random vector is still normal.

**Theorem.** Suppose $X = (X_1, X_2, \ldots, X_d)'$ has normal distribution $N(\mu, \Sigma)$. Then for any $m \times d$ matrix $C$ and $m \times 1$ vector $b$, the linear transform

$$Y = CX + b$$

is still normal with distribution $N(C\mu + b, C\Sigma C')$. 
4. Suppose \( X \sim N(\mu, \Sigma) \) and \( Y \sim N(\bar{\mu}, \bar{\Sigma}) \). Assume \( X \) and \( Y \) are independent.

Then \( X + Y \) has distribution

\[
N(\mu + \bar{\mu}, \Sigma + \bar{\Sigma}).
\]
5. **Theorem:** Suppose $X \sim N(\mu, \Sigma)$. Let $\theta_1$ and $\theta_2$ be two subvectors of $X$. Then $\theta_1$ and $\theta_2$ are independent if and only if 
\[ \text{Cov}[\theta_1, \theta_2] = 0. \]

**Corollary:** Suppose $X \sim N(\mu, \Sigma)$. Then $X_i$ and $X_j$ are independent if and only if $\sigma_{ij} = \text{Cov}[X_i, X_j] = 0$. 
6. Any conditional distribution is still normal.

**Theorem.** Suppose \( X = (X_1, X_2, \ldots, X_d)' \sim N(\mu, \Sigma) \). Suppose we partition \( X \) into subvectors

\[
X = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix},
\]

and denote

\[
\mu_1 = E[\theta_1], \quad \mu_2 = E[\theta_2], \\
\Sigma_{11} = \text{Var}[\theta_1], \quad \Sigma_{22} = \text{Var}[\theta_2], \\
\Sigma_{12} = \text{Cov}[\theta_1, \theta_2], \quad \Sigma_{21} = \text{Cov}[\theta_2, \theta_1] = \Sigma'_{12}.
\]

Then the conditional distribution of \( \theta_1 \), given \( \theta_2 \), is \( N(\bar{\mu}, \bar{\Sigma}) \) with

\[
\bar{\mu} = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (\theta_2 - \mu_2) \\
\bar{\Sigma} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}
\]
**How does the proof go?**

Suppose $X \sim N(\mu, \Sigma)$. Then for any $t = (t_1, t_2, \ldots, t_d)'$

$$M(t) = E \left[ e^{\langle t, X \rangle} \right] = e^{\langle t, \mu \rangle} + \frac{1}{2} t' \Sigma t$$
1. Suppose two different tests $A$ and $B$ are to be given a randomly chosen student. Assume that the population mean score on test $A$ is 85 with standard deviation 10, that the mean score on test $B$ is 90 with standard deviation 16, that the scores on the two test have a bivariate normal distribution with correlation 0.8. Find the probability that the student’s score on test $A$ will be greater than his score on test $B$. 
2. Suppose two random variables $X$ and $Y$ has bivariate normal distribution and

$$\text{Var}[X] = \text{Var}[Y].$$

Argue that $X + Y$ and $X - Y$ are independent.
**How we simulate $N(\mu, \Sigma)$?**

Find a matrix $A$ such that

$$\Sigma = AA'$$

Simulate $Y_1, Y_2, \ldots, Y_d$ iid $N(0, 1)$. Write $Y = (Y_1, Y_2, \ldots, Y_d)'$. Then

$$X = AY + \mu$$

has distribution $N(\mu, \Sigma)$. 

The $t$ distribution

It is also known as Student’s $t$-distribution in honor of W.S. Gosset, who publish his studies of this distribution in 1908 under pen-name “Student”.

**Definition:** Suppose $Z \sim N(0, 1)$ and $W \sim \chi^2(k)$. If $Z$ and $W$ are independent, then

$$T \doteq \frac{Z}{\sqrt{W/k}}$$

is said to have $t$-distribution with $k$ degrees of freedom (d.f.), denoted by $t(k)$. 
The density for the $t$-distribution with $k$ degrees of freedom is

$$f(x) = C \left( 1 + \frac{x^2}{k} \right)^{-\frac{k+1}{2}}.$$

Symmetric, heavier tails than normal.

When $k \to \infty$, $t(k) \to N(0, 1)$. 
Distributions associated with Normal samples

Assume that $X_1, X_2, \ldots, X_n$ are iid samples from $N(\mu, \sigma^2)$. Then

1. Sample variance

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

is independent of sample mean $\bar{X}$, and

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

2. The ratio:

$$\frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1)$$

What happens for the ratio when $S$ is replaced by $\sigma$?
A general proof use the concept of orthogonal matrix \((C' C = I)\) to show that \(\bar{X}\) and \(S\) are independent. The special case of 2-D is rather simple, and can be found in the textbook [Theorem 7.3].
APPLICATIONS OF t-DISTRIBUTION

SMALL SAMPLE ESTIMATION:

1. Normal population mean. $X_1, X_2, \ldots, X_n$ iid samples from $N(\mu, \sigma^2)$. Both $\mu$ and $\sigma$ unknown.

The $1 - \alpha$ confidence interval for $\mu$ is

$$\bar{X} \pm t_{\alpha/2}(n - 1) \frac{S}{\sqrt{n}}$$
2. Difference of means of two normal populations.

\( X_1, X_2, \ldots, X_n \) iid samples from Population 1 of \( N(\mu_1, \sigma_1^2) \).

\( Y_1, Y_2, \ldots, Y_m \) iid samples from Population 2 of \( N(\mu_2, \sigma_2^2) \).

Assumption: \( \sigma_1^2 = \sigma_2^2 = \sigma^2 \).

The \( 1 - \alpha \) confidence interval for \( \mu_1 - \mu_2 \) is

\[
(\bar{X} - \bar{Y}) \pm t_{\alpha/2}(n + m - 2)S_p\sqrt{\frac{1}{n} + \frac{1}{m}}
\]

where \( S_p \) is the pooled estimate for \( \sigma^2 \):

\[
S_p^2 = \frac{(n - 1)S_x^2 + (m - 1)S_y^2}{n + m - 2}
\]
**Small Sample Hypothesis testing:**

1. Normal population mean. \( X_1, X_2, \ldots, X_n \) iid samples from \( N(\mu, \sigma^2) \). Both \( \mu \) and \( \sigma \) unknown.

\[
H_0 : \mu = \mu_0, \quad H_a : \mu \neq \mu_0
\]

\[
P\text{-value} = P(t(n-1) > |T|)
\]

where

\[
T = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}}
\]
2. Difference of means of two normal populations.

   \( X_1, X_2, \ldots, X_n \) iid samples from Population 1 of \( N(\mu_1, \sigma_1^2) \).

   \( Y_1, Y_2, \ldots, Y_m \) iid samples from Population 2 of \( N(\mu_2, \sigma_2^2) \).

   Assumption: \( \sigma_1^2 = \sigma_2^2 = \sigma^2 \).

   \( H_0 : \mu_1 - \mu_2 = D_0, \quad H_a : \mu_1 - \mu_2 \neq D_0 \)

   \( P\)-value = \( P(t(n + m - 2) > |T|) \)

   where

   \[
   T = \frac{\bar{X} - \bar{Y} - D_0}{S_p \sqrt{\frac{1}{n} + \frac{1}{m}}}
   \]
EXAMPLES

1. Comparison of weight gain by two lots of rats under two diets.

<table>
<thead>
<tr>
<th>Diet</th>
<th>n</th>
<th>sample mean gain $\bar{x}$ (g)</th>
<th>$\sum_i(x_i - \bar{x})^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Protein</td>
<td>12</td>
<td>120</td>
<td>5032</td>
</tr>
<tr>
<td>Low Protein</td>
<td>7</td>
<td>101</td>
<td>2552</td>
</tr>
</tbody>
</table>

Do these two diets yield different weight gain?

*Solution:* $df = 12 + 7 - 2 = 17$. Pooled sample variance

$$S_p^2 = \frac{5032 + 2552}{12 + 7 - 2} = 446.12.$$  

$$T = \frac{120 - 101 - 0}{S_p\sqrt{\frac{1}{12} + \frac{1}{7}}} = 1.89.$$  

$P$-value = $P(t(17) > |1.89|) = 0.08$

Discussion on the assumptions.
Handout of normal plot.
The $F$-distribution

Suppose $X \sim \chi^2(n)$ and $Y \sim \chi^2(m)$, and that $X$ and $Y$ are independent. Then

$$W = \frac{X/n}{Y/m}$$

is said to have $F$ distribution with $n$ degrees of freedom for numerator and $m$ degrees of freedom for denominator, denoted by $F(n, m)$.

The density of $F(n, m)$ is

$$g(x) = C \cdot x^{\frac{n}{2} - 1} \left[ x + \frac{m}{n} \right]^{-\frac{m+n}{2}}, \quad x \geq 0.$$