STA260 Notes
By sytez

Preface

These notes were made for the course STA260: Probability and Statistics II at the University of Toronto Mississauga. They are based on the lectures and lecture slides of the 2024 summer offering. This course was instructed by Professor Luai Al Labadi (The Goat). These notes primarily consist of a summary of the defintions and theorems from the course. The notes are not exhaustive and may contain errors. The proofs provided are not complete and instead provide a brief outline of the proof containing the main ideas. Theorems that do not contain a proof are usually trivial to prove with exceptions of the ones whose proof is beyond the scope of this course.

Contents

7 Chapter 7 — Sampling Distributions and the Central Limit Theorem

7.1 Introduction

Definition 7.1.1. A **population** consists of the entire collection of the observations with which we are concerned.

Definition 7.1.2. A **sample** is a subset of a population.

Definition 7.1.3. A **parameter** is a numerical summary of a population. For example, the mean, the variance, etc. In practice, it is unknown.

Definition 7.1.4. A **statistic** is a numerical summary of a sample. For example, the sample mean, the sample variance, etc.

Definition 7.1.5. The statistic varies from sample to sample and hence it is a random variable and has a probability distribution called the **sampling distribution**.

The knowledge of the sampling distribution of a statistic helps to make an inference about the corresponding population (true) parameter.

Definition 7.1.6. We say that the random variables $Y_1, Y_2, \ldots, Y_n$ are **independent and identically distributed** (i.i.d.) if they are independent random variables and have the same probability distribution (same pdf/cdf).

For a random sample, we write

$$Y_1, \ldots, Y_n \overset{\text{i.i.d.}}{\sim} f(y) \quad \text{(continuous)}$$

$$Y_1, \ldots, Y_n \overset{\text{i.i.d.}}{\sim} p(y) \quad \text{(discrete)}$$

7.2   Sampling Distributions

Definition 7.2.1. The **sample mean** is defined as

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Definition 7.2.2. The **sample variance** is defined as

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$$

Theorem 7.2.1. If $Y_1, \ldots, Y_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$, then

$$\overline{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Proof. Show MGF of $\overline{Y}$ is the MGF of $N\left(\mu, \frac{\sigma^2}{n}\right)$ then conclude by uniqueness of MGF.
$\square$

Definition 7.2.3. The **standard normal distribution** is defined as

$$Z = \frac{Y - \mu}{\sigma} = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} = N(0,1)$$

Theorem 7.2.2. Let $Y_1, \ldots, Y_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$. If $Z_i = \frac{Y_i - \mu}{\sigma}$, then

$$\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n}\left(\frac{Y_i - \mu}{\sigma}\right)^2$$

has a χ^2 distribution with n degrees of freedom (df)

Proof. Note 3 facts:

1. $\chi^2_{(v)} = \text{Gamma}(v/2, 2)$

2. $\chi^2_{(v_1)} + \chi^2_{(v_2)} \sim \chi^2_{(v_1 + v_2)}$ assuming independence

3. $[N(0,1)]^2 = Z^2 \sim \chi^2_{(1)}$

The proof becomes trivial from here.
$\square$

Note that similarly we have the sum of many distributions can be studied easily.

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{Ber}(p) \implies \sum_{i=1}^{n} Y_i \sim \text{Bin}(n, p)$$

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{Poisson}(\lambda) \implies \sum_{i=1}^{n} Y_i \sim \text{Poisson}(n\lambda)$$

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \chi^2(v_i) \implies \sum_{i=1}^{n} Y_i \sim \chi^2_{\left(\sum_{i=1}^{n} v_i\right)}$$

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} N(\mu_i, \sigma_i^2) \implies \sum_{i=1}^{n} Y_i \sim N\left(\sum_{i=1}^{n}\mu_i, \sum_{i=1}^{n}\sigma_i^2\right)$$

Proof. Note the following fact:

Let $U = Y_1 + \ldots + Y_n$ then we have $M_U(t) = M_{Y_1}(t) \cdot \ldots \cdot M_{Y_n}(t)$.

That is that the MGF of the sum of random variables is the product of the MGFs of the random variables.

The proof follows trivially using the unqiueness of MGFs. $\qquad\square$

Theorem 7.2.3. As a corollary of theorem 1.2.2

If $Y_1, \ldots, Y_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$, then

$$U = \sum_{i=1}^{n} \left(\frac{Y_i - \mu}{\sigma} \right)^2 \sim \chi^2_{(n)}$$

similarly one sees that if $Y_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$ (and independent), then

$$U = \sum_{i=1}^{n} \left(\frac{Y_i - \mu_i}{\sigma_i} \right)^2 \sim \chi^2(n).$$

Theorem 7.2.4. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a normal distribution with mean μ and variance σ^2. Then

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \frac{(Y_i - \overline{Y})^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Additionally, $\overline{Y}$ and S^2 are independent.

7.3   The t-Distribution

Definition 7.3.1. Let $Z \sim N(0,1)$ and $W \sim \chi^2_{(v)}$. Then if Z and W are independent then we say

$$T = \frac{Z}{\sqrt{W/v}} \sim t_{(v)}$$

has a t-**Distribution** with v degrees of freedom.

So far it was assumed that the population standard deviation σ is known. However, this assumption may be unreasonable. We want to estimate both μ and σ. A natural statistic to deal with inferences on μ is

$$T = \frac{\overline{Y} - \mu}{S/\sqrt{n}} \sim t_{(n-1)}$$
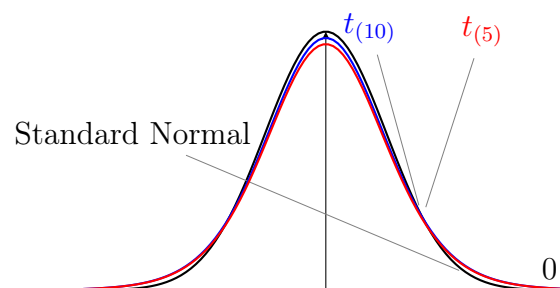
Properties of the t-Distribution

Figure 1: Comparison of Standard Normal and t-distribution curves with 5 and 10 degrees of freedom

- Each t_ν curve is bell-shaped and centered at 0.

- Each t_ν curve is spread out more than the standard normal (Z) curve.

- As ν increases, the spread of the corresponding t_ν curve decreases.

- As $\nu \to \infty$, the sequence of t_ν curves approaches the standard normal curve (the Z curve is called a t curve with df $= \infty$).

7.4   The F-Distribution

Definition 7.4.1. Let $W_1 \sim \chi^2_{(\nu_1)}$ and $W_2 \sim \chi^2_{(\nu_2)}$ be independent. Then

$$F = \frac{W_1/\nu_1}{W_2/\nu_2}$$

has a F-**Distribution** with ν_1 and ν_2 degrees of freedom.

Theorem 7.4.1. If S_1^2 and S_2^2 are the variances of independent random samples of size n_1 and n_2 taken from normal populations with variances σ_1^2 and σ_2^2, respectively, then

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1).$$

Proof. Recall that

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

the proof trivially follows. $\qquad\qquad\square$

Theorem 7.4.2. Let Y be a random variable such that $Y \sim F_{(v_1, v_2)}$. Then

$$\frac{1}{Y} \sim F_{(v_2, v_1)}$$

Theorem 7.4.3. Let $Y_1 \sim F_{(v_1, v_2)}$ then

$$\left(1 + \frac{v_1}{v_2} Y_1\right)^{-1} \sim \text{Beta}\left(\frac{v_2}{2}, \frac{v_1}{2}\right)$$

Proof. Recall that

$$\chi^2_{(v)} = \text{Gamma}\left(\frac{v}{2}, 2\right) \quad \text{and} \quad \frac{\text{Gamma}(\alpha, \gamma)}{\text{Gamma}(\alpha, \gamma) + \text{Gamma}(\beta, \gamma)} = \text{Beta}(\alpha, \beta)$$

the proof follows trivially using the defintion of the F-distribution. $\qquad\square$

Theorem 7.4.4. Let T be a random variable with a t-distribution with v degrees of freedom. Then $U = T^2$ has an F-distribution with 1 and v degrees of freedom.

7.5   The Central Limit Theorem