# Lec 2. Sampling distributions of $\bar{X}_n$ and $S_n^2$

Zhang W.P.

# Contents

1.1 $\chi^2$, $t$ and $F$ distributions . . . . . 1
1.2 Normal population case . . . . . 4
1.3 Asymptotic results for non-normal case . . . . . 6

## Learning objectives

After this section, you should be able to ...

- understand the distributions of $\bar{X}_n$ and $S_n^2$
- understand the delta theorem
- understand the orthogonal transformations of multivariate Normal vectors

## 1.1 $\chi^2$, $t$ and $F$ distributions

### Chi-square distribution

Def. If $X_1, X_2, \cdots, X_n$ i.i.d. $\sim N(0,1)$, then $X = \sum_{i=1}^{n} X_i^2$ is distributed according to the chi-square distribution with $n$ degrees of freedom. This is usually denoted as $X \sim \chi_n^2$.

The probability density function (pdf) of $\chi_n^2$ is

$$g_n(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{1.1}$$

## Student $t$ distribution

If $X \sim N(0,1), Y \sim \chi^2_n$, and $X$ is independent of $Y$, then

$$T = \frac{X}{\sqrt{Y/n}}$$

Def.

is distributed as the *student* t distribution with $n$ degrees of freedom. This is usually denoted as $T \sim t_n$.

The probability density function (pdf) of $t_n$ is

$$t_n(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < x < \infty \tag{1.2}$$

## F distribution

If $X \sim \chi^2_m$, $Y \sim \chi^2_n$, and $X$ is independent of $Y$, then

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

Def.

has $F$ distribution with $n$ and $m$ degrees of freedom. This is usually denoted as $F \sim F_{m,n}$.

If a random variable $F \sim F_{m,n}$, then the pdf of $F$ is given by

$$f_{m,n}(x) = \begin{cases} \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})} m^{\frac{m}{2}} n^{\frac{n}{2}} x^{\frac{m}{2}-1} (n+mx)^{-\frac{m+n}{2}}, & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{1.3}$$

## 1.2 Normal population case

**Theorem 1.** *If $X_1, \cdots, X_n$ are independent random variables with $X_k \sim N(a_k, \sigma_k^2)$, for $k = 1, \ldots, n$, and $c_1, c_2, \cdots, c_n$ are constants not all zero, then*

$$T = \sum_{k=1}^{n} c_k X_k \sim N\left( \sum_{k=1}^{n} c_k a_k, \sum_{k=1}^{n} c_k^2 \sigma_k^2 \right)$$

*In particular, when $c_1 = \cdots = c_n = 1/n, a_1 = \ldots = a_n, \sigma_1^2 = \ldots = \sigma_n^2$, that is, $T = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}$, we have*

$$\bar{X} \sim N(a, \sigma^2/n).$$

*Proof.* It can be proved easily by the reproducibility of normal distribution, or using the characteristic functions. □

**Theorem 2.** *If* $X_1, X_2, \cdots, X_n$ *i.i.d.* $\sim N(a, \sigma^2)$, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ *and* $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ *are sample mean and sample variance, respectively, then*

(1) $\bar{X} \sim N(a, \frac{1}{n}\sigma^2)$;
(2) $(n-1)S_n^2/\sigma^2 \sim \chi^2_{n-1}$;
(3) $\bar{X}$ *and* $S_n^2$ *are independent.*
(4) $T_n = \frac{\sqrt{n}(\bar{X}-a)}{S_n} \sim t_{n-1}$.

Hints: orthogonal transformations of independent Normal variables.

## 1.3 Asymptotic results for non-normal case

**Theorem 3.** *(Central Limit Theorem, CLT). Let $X_1, \ldots, X_n$ be an iid sample from distribution with mean $\mu$ and variance $\sigma^2 < \infty$. For $\bar{X}_n$ the sample mean, let $Z_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma$. Then $Z_n$ converges in distribution to $N(0,1)$ as $n \to \infty$.*

**Theorem 4.** *(Delta Theorem). For a sequence* $\{T_n\}$*, suppose there are numbers* $\theta$ *and* $v_\theta > 0$ *such that* $\sqrt{n}\,(T_n - \theta)$ *converges in distribution to* $\mathrm{N}\,(0, v_\theta)$*. Let* $g(t)$ *be a function, differentiable at* $t = \theta$*, with* $g'(\theta) \neq 0$*. Then* $\sqrt{n}\,[g\,(T_n) - g(\theta)]$ *converges in distribution to* $\mathrm{N}\left(0, [g'(\theta)]^2\, v_\theta\right)$*, as* $n \to \infty$*.*

**Theorem 5.** *(Central Limit Theorem, CLT). Let $X_1, \ldots, X_n$ be an iid sample from distribution with mean $\mu$ and variance $\sigma^2 < \infty$. For $S_n^2$ the sample variance, we have*

*(1)* $S_n^2 \xrightarrow{P} \sigma^2$

*(2)* $\sqrt{n}(S_n^2 - \sigma^2) \xrightarrow{\mathcal{L}} N(0, \mu_4 - \sigma^4)$, *where* $\mu_4 = E(X_1 - \mu)^4$.

**Theorem 6.** *Let $X_1, \ldots, X_n$ be an iid sample from distribution with mean $\mu$ and variance $\sigma^2 < \infty$. For $\bar{X}_n$ the sample mean and $S_n^2$ the sample variance, let $T_n = \frac{\sqrt{n}(\bar{X}-\mu)}{S_n}$, then $T_n \xrightarrow{\mathcal{L}} N(0,1)$.*

Let $X_1, \ldots, X_n$ be an independent sample from Exp($\theta$) whose mean is $\theta$. For sample mean $\bar{X}$, find the asymptotic distribution of $\bar{X}^{1/2}$.

↑Exam
↓Exam

It follows from the CLT that $\sqrt{n}(\bar{X} - \theta) \to \mathrm{N}\left(0, \theta^2\right)$ in distribution. Now consider $\bar{X}^{1/2}$. In the Delta Theorem context, $T_n = \bar{X}$ and $g(t) = t^{1/2}$. By simple calculus, $g'(\theta) = 1/2\theta^{-1/2}$. Therefore, by the Delta Theorem, $\sqrt{n}(\bar{X}^{1/2} - \theta^{1/2}) \to \mathrm{N}(0, \theta/4)$ in distribution.