Mathematical Statistics I

Chapter 6: Distributions Derived from the Normal Distribution

Jesse Wheeler

Outline

1. χ^2 distributions

2. The t and F distributions

3. Sampling Distributions

χ^2 distributions

Introduction

- This material comes primarily from Rice (2007, Chapter 6).
- Here, we introduce several important distributions that arise from transformations applied to normal distributions.
- Many of these distributions form the basis of traditional statistical inference procedures that are taught in introductory statistics courses.
- They are very useful in practice due to the central limit theorem: with enough observations, the limiting behavior of nearly all distributions is normal, so distributions that come from the normal distribution arise in practice as well.

χ^2_ν Distribution

- The first distribution we will consider is the χ^2_1 (Chi-square with 1 degree of freedom).

Definition: χ^2_1 distribution

If Z is a standard normal random variable, then $X = Z^2$ is called the chi-square distribution with 1 degree of freedom.

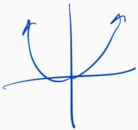
- We typically use the notation $X \sim \chi^2_1$ (in LaTeX: `\chi`).

The pdf of χ^2_1

Let X follow a χ^2_1 distribution. Then, the pdf of X is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}} x^{-1/2} e^{-x/2}. \propto x^{-\alpha} e^{-\lambda x}$$

$Z \sim N(0,1)$, $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$



$X \overset{d}{=} Z^2$

$$\begin{cases} \text{"cdf method"} \star \\ \text{change of variables: } f_X(x) = f_Z(g^{-1}(x)) \cdot \left|\frac{d}{dx} g^{-1}(x)\right| \end{cases}$$

χ^2_ν Distribution III

- In Chapter 2, we previously noted that that $f_X(x)$ is an example of a Gamma distribution.
- Specifically, the *kernel* of the Gamma density is x raised to some power, and e raised to some multiple of x:

$$f_{\text{Gamma}}(x) \propto x^{\alpha-1} e^{-\lambda x}.$$

$\rightarrow \int_0^\infty x^{\alpha-1} e^{-\lambda x} dx$

$\rightarrow f_{gamma}(x) = C(\alpha, \lambda) x^{\alpha-1} e^{-\lambda x}$

- Thus, ignoring the constant for a moment, if $\alpha = 1/2$, $\lambda = 1/2$, then the pdf of $X \sim \chi^2_1$ is just this Gamma density:

$$f_X(x) \propto x^{-1/2} e^{-x/2} = x^{\alpha-1} e^{-\lambda x}.$$

- Since both functions are proper probability density functions, they have to integrate to one, so the normalizing constant *must* be the same.

$f(x|\theta)$ use data that come from $f(x;\theta)$ to estimate θ

θ is treated as a R.V.

$$\pi(\theta|X) = \frac{f(x|\theta)\,\pi_0(\theta)}{\underbrace{\int f(x|\theta)\,\pi_0(\theta)\,d\theta}}$$

obs data

is only a function of X, not θ.

$$\pi(\theta|x) = \underline{c(x)}\, f(x|\theta)\,\pi_0(\theta)$$

$$\pi(\theta | x) \propto f(x | \theta)\, \pi_0(\theta)$$

χ^2_ν Distribution IV

- This is also easily verified. The normalizing constant of the Gamma distribution is $\lambda^\alpha/\Gamma(\alpha)$.
- With our specific values of $\lambda = \alpha = 1/2$, and recalling that $\Gamma(1/2) = \sqrt{\pi}$,

$$\frac{1}{\sqrt{2\pi}} = \frac{(1/2)^{(1/2)}}{\Gamma(1/2)} = \frac{\lambda^\alpha}{\Gamma(\alpha)}$$

MGF of χ^2_1

We previously derived the MGF of a Gamma(α, λ) distribution: $M(t) = \big(\lambda/(\lambda - t)\big)^\alpha$. Thus, the MGF of a Chi-square(1) distribution is

$$M(t) = (1 - 2t)^{-1/2}, \quad t < 1/2.$$

χ^2_ν Distribution V

Definition

If $U_1, U_2, \ldots, U_n$ are n independent χ^2_1 random variables, then

$$V = U_1 + U_2 + \ldots + U_n$$

then the distribution of V is called the Chi-square distribution with n degrees of freedom, denoted χ^2_n.

- There are a few different ways of deriving the pdf of a χ^2_n random variable. Here, we will use the MGF uniqueness theorem.

$V = U_1 + U_2 + \cdots + U_n$

- Let $M_i(t)$ denote the MGF of U_i, where $U_i \sim \chi^2_1$. Then, due to independence,

$$M_V(t) = M_{\sum_i U_i}(t) = \prod_{i=1}^{n} M_i(t) = \big(M_1(t)\big)^n = (1-2t)^{-n/2}$$

$$= \left(\frac{1}{1-2t}\right)^{n/2}$$

- Compare this to the Gamma MGF: $\boxed{M(t) = \big(\lambda/(\lambda - t)\big)^\alpha}$. Then, setting $\underline{\lambda = 1/2}$, $\underline{\alpha = n/2}$, we see that $\underline{V}$ has a $\underline{\text{Gamma}(n/2, 1/2)}$ distribution. $V \sim \chi^2_n$
- Thus, the pdf of V is given by:

$$f_V(v) = \frac{1}{2^{n/2}\Gamma(n/2)} v^{(n/2)-1} e^{-v/2}.$$

χ^2_ν Distribution VII

- The expected value and variance of the χ^2_n distribution can easily be found then by using the fact that it is a special case of a Gamma distribution.

The t and F distributions

The Student's t distributions

The Student's t distribution

If $Z \sim N(0,1)$ and $U \sim \chi_n^2$, and Z and U are independent, then the distribution of T, where

$$T = \frac{Z}{\sqrt{U/n}},$$

is called the Student's t distribution (or simply the t distribution) with n degrees of freedom, which is often denoted t_n

- Students often forget the make sure that Z and U in the definition of the t distribution are independent.
- The t distribution is the distribution used to perform the famed "t-test".

The density of the t_n distribution

The pdf of the t distribution with n degrees of freedom is:

$$f(t) = \frac{\Gamma\big((n+1)/2\big)}{\sqrt{n\pi}\Gamma(n/2)}\Big(1 + \frac{t^2}{n}\Big)^{-(n+1)/2}$$

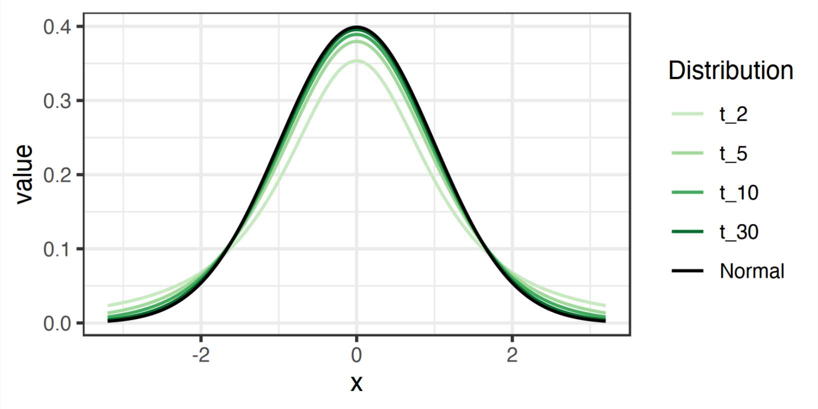
$P(T \leq t) = P\left(\frac{Z}{\sqrt{V/n}} \leq t\right)$

→ CDF, Quotient of 2 ind. RV. $\frac{X}{Y}$, Jacobian

- The derivation of the pdf of a t distribution is a good practice exercise.
- Recall it is defined as the ratio of two independent random variables; in Chapter 3, we derived a formula for computing densities of random variables of this form.
- Note that $f(t) = f(-t)$, and so f is symmetric about zero.
- It also has a bell-curve shape similar to a normal distribution.

The Student's t distributions III

- You can see as $n \to \infty$, the t_n distribution converges to the standard normal (e.g., use Slutsky's theorem, good practice).



The F distributions

Sampling Distributions

The sample mean

- In what follows, we'll assume that we are taking samples $X_1, X_2, \ldots, X_n$ from a larger population.
- These samples can be repeated experiments, or repeated observations. However, we will assume in general that the samples are independent and identically distributed, unless stated otherwise.
- For the remainder of the chapter, we will also assume $X_i \sim N(\mu, \sigma^2)$ for all i.

The sample mean II

- As a reminder from earlier chapters, linear combinations of independent normal random variables are also normally distributed. Thus, if $X_1, X_2, \ldots, X_n$ are iid normal, then $\bar{X}_n$ is also normally distributed. $\frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n = \frac{1}{n}\sum X_i$

Sampling distribution of the mean

If X_i are iid $N(\mu, \sigma^2)$, then $\bar{X}_n$ is normal, with

$$E\left[1/n \sum_i X_i\right] = (1/n)\sum_i \mu = \mu,$$

$$\mathrm{Var}\left(1/n \sum_i X_i\right) = 1/n^2 \sum_i \sigma^2 = \sigma^2/n.$$

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

- Sampling distribution of $\bar{X}_n$ is $\overbrace{N(\mu, \frac{\sigma^2}{n})}^{\text{exact.}}$.

Y is population, $\bar{Y}_n \approx N(\mu, \frac{\sigma^2}{n})$

The sample mean III

Lemma 6.1: Independent Normal RVs

Let X and Y be normally distributed random variables. Then X and Y are independent, if and only if

$$\mathrm{Cov}(X, Y) = 0.$$

If independent, the Cov(x,y) = 0.

- The above statement can be proved using the factorization theorem, and considering the MGF or pdf of a bivariate normal distribution.
- Recall that for most distributions, independence implies $\mathrm{Cov}(X, Y) = 0$, but not the other way around.
- It turns out that the normal distribution is the only distribution that has this property.

The sample mean IV

Theorem 6.1: Independence of Deviations

Let $X_1, \ldots, X_n$ be iid $N(\mu, \sigma^2)$ random variables. Then, $\bar{X}_n$ is independent of the vector of random variables called the *deviations*, $(X_i - \bar{X}_n)_{i=1}^n$.

sample average

Proof.

$D_i = (X_i - \bar{X}_n)$

$\underline{\bar{X}_n} \perp \underline{X_i - \bar{X}_n}$

(independence).

- $\bar{X}_n$ is Normal$(\mu, \frac{\sigma^2}{n})$
- $X_i - \bar{X}_n$ is normal

cov is zero implies independence.

$$\operatorname{cov}(\bar{X}_n, X_i - \bar{X}_n) = \operatorname{cov}(\bar{X}_n, X_i) - \operatorname{cov}(\bar{X}_n, \bar{X}_n)$$

$$= \operatorname{cov}\left(\frac{1}{n}\sum_{j=1}^{n} X_j, X_i\right) - \operatorname{cov}\left(\frac{1}{n}\sum_{j=1}^{n} X_j, \frac{1}{n}\sum_{k=1}^{n} X_k\right)$$

$$= \frac{1}{n}\operatorname{cov}(X_i, X_i) - \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\operatorname{cov}(X_i, X_j)$$

$$= \frac{1}{n}\underbrace{\operatorname{cov}(X_i, X_i)}_{} - \frac{1}{n^2}\sum_{j=1}^{n}\operatorname{cov}(X_j, X_j)$$

$$= \frac{1}{n}\sigma^2 - \frac{1}{n^2}\sum_{j=1}^{n}\sigma^2$$

$$= \frac{1}{n}\sigma^2 - \frac{1}{n}\sigma^2 = \boxed{0}$$

$\leftarrow$ $\bar{X}_n$ is ind. of $X_i - \bar{X}_n$

Corollary 6.1

If the X_i are iid $N(\mu, \sigma^2)$, then $\bar{X}_n$ is independent of the sample variance S^2, defined by

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

$\bar{X}_n \perp S^2$

a function of $(X_i - \bar{X}_n)$

Theorem 6.2

If the X_i are iid normal, then $(n-1)S^2/\sigma^2$ has a chi-square distribution with $n-1$ degrees of freedom.

Proof.

$$\left(\frac{X_i-\mu}{\sigma}\right) \sim N(0,1)$$

$$\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2_1$$

$$\frac{1}{\sigma^2}\sum (X_i-\mu)^2 = \sum \left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2_n$$

$$\frac{1}{n}\sum X_i - \bar{X}_n \to 0$$

$$* \quad \underbrace{\frac{1}{\sigma^2}\sum (x_i-\mu)^2}_{\chi^2_n} = \frac{1}{\sigma^2}\sum\left[(x_i-\bar{x}_n)+(\bar{x}_n-\mu)\right]^2$$

$$= \underbrace{\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2}_{(n-1)S^2/\sigma^2} + \underbrace{\left(\frac{\bar{x}_n-\mu}{\sigma/\sqrt{n}}\right)^2}_{U\sim\chi^2_1}$$

$V\sim\chi^2_n$

$$V = (n-1)S^2/\sigma^2 + U \qquad V\sim\chi^2_n, \quad U\sim\chi^2_1$$

$$(n-1)S^2/\sigma^2 = V-U$$

$$M_{(n-1)S^2/\sigma^2}(t) = M_{V-U}(t) = \frac{M_V(t)}{M_U(t)} = \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}}$$

$$= (1-2t)^{-(n-1)/2}$$

mgf of χ^2_{n-1}

The sample mean VII

Theorem 6.3: t distribution

Let X_i be iid Normal(μ, σ^2) random variables, and let $\bar{X}_n$ and S^2 denote the sample mean and variance, respectively. Then,

$$\frac{\bar{X}_n - \mu}{S/\sqrt{n}} \sim t_{n-1}.$$

t-distn defined as ratio of $N(0,1)$, and sqrt of χ^2_n / n : $\dfrac{Z}{\sqrt{V/n}}$

proof:

$$\frac{\bar{X}_n - \mu}{S/\sqrt{n}} = \frac{\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right)}{\sqrt{S^2/\sigma^2}} \sim t_{n-1}.$$

$\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right) \to N(0,1)$

$\sqrt{S^2/\sigma^2} \sim \chi^2_{n-1}$

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

Final Comments

- The identity in Theorem 6.3 provides the theoretical justification for a one-sample t-test. The justification for a two-sample t-test is derived in a similar way.
- Theorem 6.3 relies on the X_i coming from a normal population, then this distribution is exact. The normal distribution arises in a large number of real-world applications.
- In practice, however, the t-test is often used even when the samples X_i do not come from a normal distribution.
- The justification of this practice can heuristically be justified by the central limit theorem: Even if the X_i are not normally distributed, $\bar{X}_n - \mu$ will be approximately normally distributed, with even relatively small n.

Final Comments II

- Note, however, that the limiting distribution of the statistic in Theorem 6.3 is normal, not t. This can be shown with Slutsky's Theorem, noting that $S_n^2 \xrightarrow{p} \sigma^2$ (see CLT example in Chapter 5 slides.)
- In practice, the t distribution is often used because it has heavier tails than the normal distribution, and thus leads to conservative estimates of the distribution of the statistic in Theorem 6.3. Simulation studies suggest this approximation is quite good, even when there are large deviations from normality.

References and Acknowledgements

Rice JA (2007). *Mathematical statistics and data analysis*, volume 371. 3 edition. Thomson/Brooks/Cole Belmont, CA.

- Compiled on November 25, 2025 using R version 4.5.1.
-
- We acknowledge students and instructors for previous versions of this course / slides.