Mathematical Statistics I

Chapter 6: Distributions Derived from the Normal Distribution

Jesse Wheeler

Contents

1 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

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

Proof. There are a few ways to show this is the case, and was one of the early examples we saw in Chapter 2. For practice, we repeat this example here.

- By definition, X has the same distribution of Z^2, where Z is a standard normal.
- Recall the standard normal density is:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

- Using the CDF method, we write

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(Z^2 \leq x) \\ &= P(-\sqrt{x} \leq Z \leq \sqrt{x}) \\ &= P(Z \leq \sqrt{x}) - P(Z \leq -\sqrt{x}) \\ &= \Phi(\sqrt{x}) - \Phi(-\sqrt{x}), \end{aligned}$$

- Where $\Phi(z)$ is the cdf of Z.
- Taking the derivative of both sides of the equation, the chain rule gives us

$$\begin{aligned} f_X(x) &= \frac{1}{2} x^{-1/2} \phi(\sqrt{x}) + \frac{1}{2} x^{-1/2} \phi(-\sqrt{x}) \\ &= x^{-1/2} \phi(\sqrt{x}), \end{aligned}$$

- where the last step is a result of the symmetry of $\phi(x)$, noting $\phi(-x) = \phi(x)$ for all $x \in \mathbb{R}$.
- Thus, replacing $\phi(\sqrt{x})$ with the definition,

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

□

- 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:

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

- 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.
- 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

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

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

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

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

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

- 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.

2 The t and F distributions

The Student's t distributions

The Student's t distribution
If $Z \sim N(0, 1)$ and $U \sim \chi^2_n$, and Z and U are independent, then the distribution of T, where

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

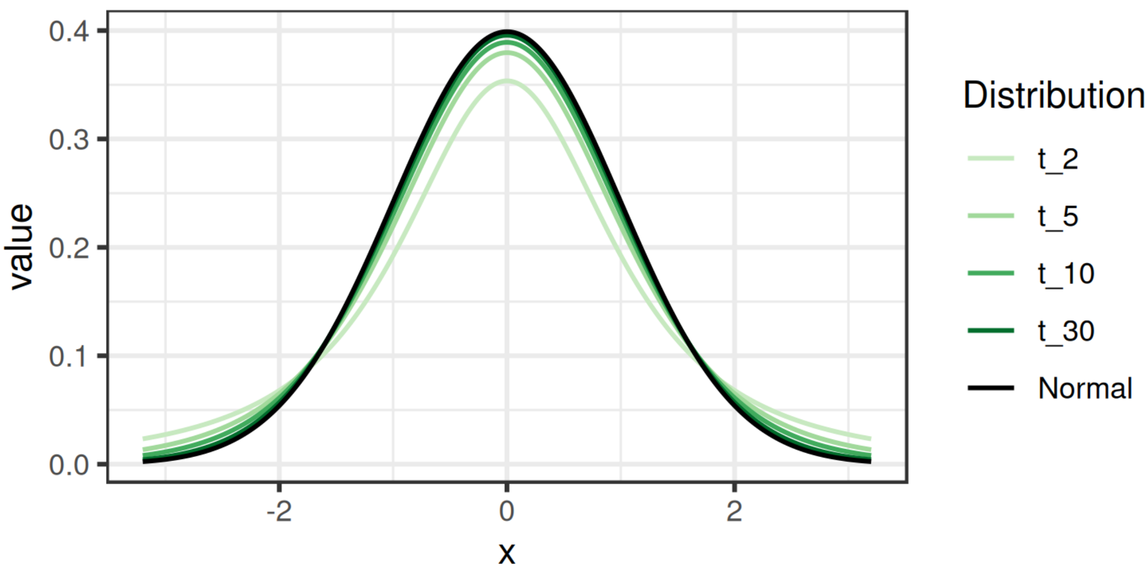
is called the Student's t distribution (or simply the t distribution) with n degrees of freedom, which is often denoted t_n

- 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:

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

- 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.
- 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

3 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.
- 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.

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

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

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

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

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*

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

- 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.

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*, $\big(X_i - \bar{X}_n\big)_{i=1}^n$.

Proof. First, note that $\bar{X}_n$ is normally distributed. Using Lemma 6.1, all we need to do is argue that the deviations are normally distributed, and that the covariance between $\bar{X}_n$ and $X_i - \bar{X}_n$ is zero. Because $\bar{X}_n$ and X_i are both normally distributed, then for all i, $X_i - \bar{X}_n$ is just a linear combination of normally distributed random variables, and as a result is also normally distributed. Using the bilinearity of covariance,

$$\begin{aligned}
\mathrm{Cov}(\bar{X}_n, X_i - \bar{X}_n) &= \mathrm{Cov}(\bar{X}_n, X_i) - \mathrm{Cov}(\bar{X}_n, \bar{X}_n) \\
&= \mathrm{Cov}\Big(\sum_{j=1}^n \frac{1}{n} X_j, X_i\Big) - \mathrm{Cov}\Big(\sum_{i=1}^n \frac{1}{n} X_i, \sum_{j=1}^n \frac{1}{n} X_j\Big) \\
&= \frac{1}{n}\mathrm{Cov}(X_i, X_i) - \sum_{i=1}^n \sum_{j=1}^n \frac{1}{n^2}\mathrm{Cov}(X_i, X_j) \\
&= \frac{\sigma^2}{n} - \frac{1}{n^2}\sum_{i=j}^n \mathrm{Cov}(X_i, X_j) \\
&= \frac{\sigma^2}{n} - \frac{1}{n^2}(n\sigma^2) = 0.
\end{aligned}$$

Thus, by Lemma 6.1, $\bar{X}_n$ is independent of $X_i - \bar{X}_n$ for all i. □

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

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

Proof. First recall from one of our homework problems that S^2 is an unbiased estimate of σ^2: $E[S^2] = \sigma^2$ (Note that this does not require the assumption that the X_i are normally distributed, just that they are independent and identically distributed with finite variance $\mathrm{Var}(X_i) = \sigma^2 < \infty$.)
From Theorem 6.1, if the X_i are normally distributed, then $\bar{X}_n$ is independent of the vector $\big(X_i - \bar{X}_n\big)_{i=1}^n$. Because S^2 is just a function of $\big(X_i - \bar{X}_n\big)_{i=1}^n$, we can immediate conclude that $\bar{X}_n$ is independent of S^2. □

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. First recall how we defined the χ^2_n distribution. First, if we square a standard normal $Z \sim N(0,1)$, then $Z^2 \sim \chi^2_1$. Then, the sum of n χ^2_1 random variables is how we define a χ^2_n random variable, i.e., if $Z_i \overset{iid}{\sim} N(0,1)$, then $Z_1^2 + \ldots + Z_n^2 \sim \chi^2_n$. Thus, because $E[X_i] = \mu$, $\text{Var}(X_i) = \sigma^2$ for all i,

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

and therefore

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

Now we will expand this same sum by adding and subtracting $\bar{X}_n$:

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

From above, we already know this entire expression results in a random variable with a χ^2_n distribution. We can now manipulate the right hand side to get the distribution of $(n-1)S^2/\sigma$, which was our original goal. We will now expand the square, and note that $\sum_{i=1}^{n}(X_i - \bar{X}) = 0$ to simplify:

$$\begin{aligned}\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 + \left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right)^2 \\ V &= (n-1)S^2/\sigma^2 + U,\end{aligned}$$

where $U \sim \chi^2_1$, since it is the result of squaring a standard normal random variable, and $V \sim \chi^2_n$. Importantly, U is a function of $\bar{X}_n$, which was shown to be independent of S^2 in Corollary 6.1. As a result, we can use the identity for the MGF of the sum of independent random variables: $M_V(t) = M_{(n-1)S^2/\sigma^2}M_U(t)$. For the values of t for which it is defined, we can then solve for the MGF of $(n-1)S^2/\sigma^2$, which uniquely determines the distribution of the random variable. As a reminder, the MGF of a χ^2_n random variable is $M(t) = (1-2t)^{-n/2}$. Thus,

$$\begin{aligned}M_{(n-1)S^2/\sigma^2} &= M_V(t)/M_U(t) \\ &= \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} \\ &= (1-2t)^{-(n-1)/2},\end{aligned}$$

which is the MGF of a χ^2_{n-1} random variable. □

The sample variance

Acknowledgments

- Compiled on November 25, 2025 using R version 4.5.2.

- We acknowledge students and instructors for previous versions of this course / slides.

References

Rice JA (2007). *Mathematical statistics and data analysis*, volume 371. 3 edition. Thomson/Brooks/Cole Belmont, CA. 1