

# 170S Week 8 Discussion Notes

Colin Ni

February 24, 2025

This week we will discuss chi-squared and  $t$ -distributions, which we have been using in our study of confidence intervals. The hope is that this will fill in any gaps in the lectures or in your previous knowledge.

**Warmup.** Let  $n > 0$  be an integer, and consider

$$\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T,$$

the difference between the  $n \times n$  identity matrix and the matrix where every entry is  $\frac{1}{n}$ . For example, for  $n = 3$  we have

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

Describe the eigenvalues and eigenvectors.

*Solution.* This is a projection matrix because

$$\left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right) \left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right) = \left(\mathbf{I} - \frac{2}{n}\mathbf{1}\mathbf{1}^T + \frac{1}{n^2}\mathbf{1}\mathbf{1}^T\mathbf{1}\mathbf{1}^T\right) = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T,$$

so by the spectral theorem (this is real and symmetric) it is orthogonal projection onto a subspace. To determine which subspace, observe that  $\mathbf{1}$  is in the kernel and that the first  $n - 1$  vectors in the matrix are linearly independent. To summarize, there exists an orthonormal basis of eigenvectors, one with eigenvalue 0 and the others with eigenvalue 1.  $\square$

## Chi-squared distributions

**Definition.** The *chi-squared distribution with  $n$  degrees of freedom* is the distribution of the sum of squares of  $n$  independent standard normals:

$$\chi^2(n) \sim Z_1^2 + \dots + Z_n^2 \quad \text{where} \quad Z_1, \dots, Z_n \sim \text{Norm}(0, 1)$$

are independent.

This is a great starter definition (different from the one in the book) because we can easily deduce some basic properties. For example  $\chi^2(1) = \text{Norm}(0, 1)$ , and if  $\chi^2(n)$  and  $\chi^2(m)$  are independent, then  $\chi^2(n) + \chi^2(m) = \chi^2(n + m)$ . These are both theorems in the textbook. Furthermore, we can deduce that the mean is

$$\mathbb{E}(\chi^2(n)) = \mathbb{E}(Z_1^2 + \dots + Z_n^2) = \mathbb{E}(Z_1^2) + \dots + \mathbb{E}(Z_n^2) = n$$

since  $1 = \text{Var}(Z_i) = \mathbb{E}(Z_i^2) - \mathbb{E}(Z_i)^2 = \mathbb{E}(Z_i^2)$ , and the variance is

$$\text{Var}(\chi^2(n)) = \text{Var}(Z_1^2 + \dots + Z_n^2) = \text{Var}(Z_1^2) + \dots + \text{Var}(Z_n^2) = 2n$$

since  $\text{Var}(Z_i^2) = \mathbb{E}(Z_i^4) - \mathbb{E}(Z_i^2)^2 = 3 - 1 = 2$  using this MSE answer.

The key property of chi-squared distributions for statistics is the following:

**Key Property.**

$$\sum_{i=1}^n (Z_i - \bar{Z})^2 \sim \chi^2(n - 1),$$

where  $Z_1, \dots, Z_n \sim \text{Norm}(0, 1)$  are i.i.d. standard normals.

In other words, subtracting off the mean from each  $Z_i$  before squaring amounts to removing one of the degrees of freedom.

*Proof of Key Property.* We begin by vectorizing. Let  $\mathbf{z} \in \mathbb{R}^n$  denote the column vector consisting of the  $Z_1, \dots, Z_n$  so that  $\bar{Z} = \frac{1}{n} \mathbf{1}^T \mathbf{z}$ . Then

$$\begin{aligned} \sum_{i=1}^n (Z_i - \bar{Z})^2 &= \sum_{i=1}^n \left( Z_i - \frac{1}{n} \mathbf{1}^T \mathbf{z} \right)^2 \\ &= \left( \mathbf{z} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{z} \right)^T \left( \mathbf{z} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{z} \right) \\ &= \mathbf{z}^T \left( \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \right)^T \left( \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \right) \mathbf{z} \\ &= \mathbf{z}^T \left( \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \right) \mathbf{z}, \end{aligned}$$

where we are using our observation from the warmup that the matrix in the middle is a projection matrix. Moreover, in our warmup we found that the dimension of the 0-eigenspace is 1 and the dimension of the 1-eigenspace is  $k - 1$ . Thus, by the spectral theorem we can find an orthonormal basis of eigenvectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  where  $\mathbf{v}_1$  has eigenvalue 0 and the others  $\mathbf{v}_2, \dots, \mathbf{v}_n$  have eigenvalue 1.

Consider

$$\mathbf{Q} = (\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n)$$

which is orthogonal in the sense that  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$ . Recall that changing bases preserves normality (*e.g.* we discussed that  $a \cdot \text{Norm}(\mu, \sigma^2) = \text{Norm}(a\mu, a^2\sigma^2)$ ), and here we in fact have

$$\mathbf{x} = \mathbf{Q}^T \mathbf{z} \sim \text{Norm}(\mathbf{0}, \mathbf{Q}^T \mathbf{I} \mathbf{Q}) = \text{Norm}(\mathbf{0}, \mathbf{I}).$$

Thus

$$\begin{aligned} \mathbf{z}^T \left( \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \right) \mathbf{z} &= \mathbf{x}^T \mathbf{Q}^T \left( \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \right) \mathbf{Q} \mathbf{x} \\ &= \mathbf{x}^T \mathbf{Q}^T \begin{pmatrix} \mathbf{0} & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix} \mathbf{x} \\ &= \mathbf{x}^T \begin{pmatrix} \mathbf{0} & \mathbf{e}_2 & \cdots & \mathbf{e}_k \end{pmatrix} \mathbf{x} \\ &= \mathbf{x}_2^2 + \cdots + \mathbf{x}_n^2 \\ &\sim \chi^2(n-1). \end{aligned} \quad \square$$

## Student's $t$ -distributions

In short,  $t$ -distributions are generalizations of the standard normal distribution, and we use them in statistics when the true variance is known.

Let  $X_1, \dots, X_n \sim \text{Norm}(\mu, \sigma^2)$  be independent samples from a normal random variable. Assume that  $\mu$  is known.

**Motivation.** When  $\sigma^2$  is also known, then by 170A/170E we have

$$\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \sim \text{Norm}(0, 1),$$

and even if  $X$  is arbitrary, we can use the central limit theorem

$$\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \rightarrow \text{Norm}(0, 1) \quad \text{as } n \rightarrow \infty$$

to approximate this quantity.

What do we do if  $\sigma^2$  is unknown? Of course we can estimate  $\sigma^2$  using the unbiased estimator  $s^2$ , the sample variance, so now we are looking at

$$\frac{\bar{X} - \mu}{\sqrt{s^2/n}}.$$

Unfortunately, this is not normal because the  $s^2$  is itself a random variable and not just a constant like  $\sigma^2$  was. But it is a tractable thing.

**Definition.** The *Student's  $t$ -distribution with  $n - 1$  degrees of freedom* is the distribution of the quantity

$$\frac{\bar{X} - \mu}{\sqrt{s^2/n}} \sim t_{n-1}.$$

Let us explain how this can be defined in terms of a chi-squared distribution. The important point is as follows:

**Proposition.**

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

Note that the mean of the LHS is  $n-1$  because  $s^2$  is an unbiased estimator for  $\sigma^2$ , and the mean of the RHS is  $n-1$  by above.

*Proof of Proposition.* We will use the Key Property of chi-squared distributions that we discussed above for the random variables

$$Z_i = \frac{X_i - \mu}{\sigma} \sim \text{Norm}(0, 1).$$

Note that these are still independent random variables, and observe that

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^n Z_i = \frac{1}{n} \sum_{i=1}^n \frac{X_i - \mu}{\sigma} = \frac{\bar{X} - \mu}{\sigma}.$$

By the aforementioned Key Property, we have

$$\begin{aligned} (n-1)\frac{s^2}{\sigma^2} &= \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \\ &= \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{\sigma} \right)^2 \\ &= \sum_{i=1}^n (Z_i - \bar{Z})^2 \\ &\sim \chi^2(n-1), \end{aligned}$$

as desired. □

Now observe that the quantity we used to define the Student  $t$ -distribution is simply

$$t_{n-1} \sim \frac{\bar{X} - \mu}{\sqrt{s^2/n}} = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \cdot \frac{1}{\sqrt{s^2/\sigma^2}} \sim \frac{\text{Norm}(0, 1)}{\sqrt{\chi^2(n-1)/(n-1)}}.$$