Math 40 Probability and Statistical Inference

Winter 2021

Yoonsang Lee (yoonsang.lee@dartmouth.edu)

Lecture 15: Four important distributions in statistics (Chapter 4)

Multivariate normal distributions (4.1)

- Section 4.1 is an extension of the bivariate normal distributions. The only difference is that now the random vector has more than two components (thus called 'multivariate').
- As Linear Algebra is not a prerequisite of this course, I will mention only the following fact

  $$\text{If } \mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Omega), \quad \mathbf{Z} = \Omega^{-1/2}(\mathbf{X} - \boldsymbol{\mu}) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}).$$

  Here $\mathbf{I}$ is the identity matrix and $\Omega^{-1/2}$ is the inverse of the square root of the covariance matrix Ω.
- I strongly recommend you to read this section and the appendix for matrix algebra (section 10.2) at your own pace.

Chi-square distributions (4.2)

- Let $X_1, X_2, ..., X_n$ be IID from the standard normal distribution.
- We are interested in the distribution of

$$\chi^2(n) = \sum_i^n X_i^2,$$

  the square sum of $X_i's$.
- We have already considered the case $n = 1$ several times (using the idea of transformation).
- In this section, we are interested in n independent sum of X_i^2.

Chi-square distributions (4.2)

- Let $X_1, X_2, ..., X_n$ be IID from the standard normal distribution.
- We are interested in the distribution of

$$\chi^2(n) = \sum_i^n X_i^2,$$

  the square sum of $X_i's$.
- The distribution of $\chi^2(n)$ is called **Chi-square with** n **degrees of freedom**.
- This is actually a Gamma distribution with $\alpha = n/2$ and $\lambda = 1/2$.

$$f(s;n) = \frac{1}{2^{n/2}\Gamma(n/2)} s^{n/2-1} e^{-s/2}, s \geq 0$$

Chi-square distributions (4.2)

- From the independence of X_i^2's whose mean is 1 and variance 2,
$$E(S) = n.$$
and
$$Var(S) = 2n.$$
- Example 4.19 shows that the MGF of the chi-square with n dof (degrees of freedom) is
$$M(t; n) = \frac{1}{(1-2t)^{n/2}}.$$
- Using the MGF, you can also check $E(S) = n$. Additionally, $E((\chi^2(n))^2) = n(n+2)$.
- Also, from the definition of the Chi-square distribution,
$$\chi^2(n_1) + \chi^2(n_2) = \chi^2(n_1 + n_2)$$

Chi-square distributions (4.2)

- If X_i iid from $\mathcal{N}(\mu, \sigma^2)$, then (example 4.20)

$$\frac{1}{\sigma^2}\sum_i^n (X_i - \mu)^2 \sim \chi^2(n)$$

  (because $\frac{X_i - \mu}{\sigma}$ is standard normal).
- In data science, we often do not know the exact value of μ (this is something we need to estimate from data). Instead, we use the sample mean $\overline{X} = \frac{1}{n}\sum_i^n X_i$
- (Theorem 4.22)

$$\frac{1}{\sigma^2}\sum_i^n (X_i - \overline{X})^2 \sim \chi^2(n-1)$$

  a chi-square with $df = n - 1$, not $df = n$.

Expectations and variances of quadratic forms (4.2.2)

We consider a random vector $\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}, \Omega)$ of size n. If $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ fixed symmetric matrices,

- $E(\mathbf{y}'\mathbf{A}\mathbf{y}) = tr(\mathbf{A}\Omega) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$
- $Var(\mathbf{y}'\mathbf{A}\mathbf{y}) = 2tr(\mathbf{A}\Omega)^2 + 4\boldsymbol{\mu}'\mathbf{A}\Omega\mathbf{A}\boldsymbol{\mu}$.
- $Cov(\mathbf{y}, \mathbf{y}'\mathbf{A}\mathbf{y}) = 2\Omega\mathbf{A}\boldsymbol{\mu}$
- $E((\mathbf{y} - \boldsymbol{\mu})'\mathbf{A}(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})'\mathbf{B}(\mathbf{y} - \boldsymbol{\mu})) = tr(\mathbf{A}\Omega)tr(\mathbf{B}\Omega) + 2tr(\mathbf{A}\Omega\mathbf{B}\Omega)$

t-distributions (4.3)

Let $X \sim \mathcal{N}(0,1)$ and $Y \sim \chi^2(n)$. Then,

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

follows the $t-$distribution with $df = n$, which is denoted as

$$X_n \sim t(n).$$

- $E(X_n) = 0$
- $Var(X_n) = \frac{n}{n-2}$
- Check the textbook for the density function.
- As $n \to \infty$, X_n converges to the standard normal (Theorem 4.29).

t-distributions (4.3)

- Let X_i be IID from $\mathcal{N}(\mu, \sigma^2)$, $i = 1, 2, 3, ..., n$. Unfortunately, we do not know μ and σ^2.
- We estimate μ and σ^2 using the sample mean and variance

$$\overline{X} = \frac{1}{n}\sum_{i}^{n} X_i,$$

and

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i}^{n}(X_i - \overline{X})^2.$$

- (Theorem 4.31; the most important fact in this section)

$$\frac{\sqrt{n}(\overline{X} - \mu)}{\hat{\sigma}} \sim t(n-1)$$

t-distributions (4.3)

- Let X_i be IID from $\mathcal{N}(\mu, \sigma^2)$, $i = 1, 2, 3, ..., n$. We have another set of data Y_j from the same distribution $\mathcal{N}(\mu, \sigma^2)$, $j = 1, 2, ..., m$. As before, we do not know μ and σ^2.
- Let $\overline{X}$ and $\overline{Y}$ be the sample mean of the two samples

$$\overline{X} = \frac{1}{n} \sum_{i}^{n} X_i, \quad \overline{Y} = \frac{1}{m} \sum_{i}^{n} Y_i.$$

- Let

$$\hat{\sigma}^2 = \frac{1}{n+m-2} \left(\sum_{i}^{n} (X_i - \overline{X})^2 + \sum_{i}^{m} (Y_i - \overline{Y})^2 \right)$$

- (Theorem 4.33; another important fact in this section)

$$\frac{\overline{X} - \overline{Y}}{\hat{\sigma} \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t(n+m-2)$$

F-distributions (4.4)

- Let $X \sim \chi^2(m)$ and $Y \sim \chi^2(n)$ be two independent random variables.
- The distribution of $\frac{X}{Y}$ is the F-distribution with degrees of freedom m and n.
- F-distributions have applications in the analysis of variance (ANOVA).
- Check your textbook for the density.
- Mean: $\frac{n}{n-2}$, mode: $\frac{n(m-2)}{m(n+2)}$, variance: $\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$.

F-distributions (4.4)

- Let X_i be IID from $\mathcal{N}(\mu, \sigma^2)$, $i = 1, 2, 3, ..., n$. We have another set of data Y_j from the same distribution $\mathcal{N}(\mu, \sigma^2)$, $j = 1, 2, ..., m$. As before, we do not know μ and σ^2.
- Not surprisingly, we have

$$\frac{\frac{\sum_i^n (X_i - \overline{X})^2}{n-1}}{\frac{\sum_i^n (Y_j - \overline{Y})^2}{m-1}} \sim F(n-1, m-1)$$