STAT 714 fa 2023 Lec 04

Distributions of quadratic forms, Cochran's theorem, ANOVA table

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Univariate Normal distribution:

- The pdf of the Normal(μ, σ^2) distribution is given by

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right] \quad \text{for } y \in \mathbb{R}.$$

- The moment generating function is $M_Y(t) = e^{t\mu + \sigma^2 t^2/2}$ for all $t \in \mathbb{R}$.
- The pdf and cdf of the Normal$(0,1)$ distribution get special notation:

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

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad \text{for } z \in \mathbb{R}.$$

- The mgf of $Z \sim$ Normal$(0,1)$ is $M_Z(t) = e^{t^2/2}$ for all $t \in \mathbb{R}$.

Moment generating function of a random vector

The *moment generating function* of a $p \times 1$ random vector $\mathbf{y}$ is defined as

$$M_{\mathbf{y}}(\mathbf{t}) = \mathbb{E}e^{\mathbf{t}^T\mathbf{y}},$$

provided the expectation exists for all $\mathbf{t}$ such that $\|\mathbf{t}\|_\infty < h$ for some $h > 0$.

Exercise: Let $Z_1, \dots, Z_p \overset{\text{ind}}{\sim} \text{Normal}(0, 1)$ and set $\mathbf{z} = [Z_1, \dots, Z_p]^T$. Find the mgf of $\mathbf{y} = \boldsymbol{\mu} + \mathbf{A}\mathbf{z}$.

Result (Multivariate mgf results)

Let $\mathbf{y}_1, \ldots, \mathbf{y}_K$ *be* K *random vectors with mgfs* $M_{\mathbf{y}_1}, \ldots, M_{\mathbf{y}_K}$*, respectively. Then*

1. $\mathbf{y}_1 \overset{d}{=} \mathbf{y}_2$ *iff* $M_{\mathbf{y}_1}(\mathbf{t}) = M_{\mathbf{y}_2}(\mathbf{t})$ *for all* $\|\mathbf{t}\|_\infty < h$ *for some* $h > 0$.
2. $\mathbf{y}_1, \ldots, \mathbf{y}_K$ *are mutually independent iff*

$$M_{\mathbf{y}}(\mathbf{t}) = M_{\mathbf{y}_1}(\mathbf{t}_1) \times \cdots \times M_{\mathbf{y}_K}(\mathbf{t}_K)$$

for all $\mathbf{t}$ *such that* $\|\mathbf{t}\|_\infty < h$ *for some* $h > 0$*, where* $\mathbf{y} = [\mathbf{y}_1^T \ldots \mathbf{y}_K^T]^T$ *and* $\mathbf{t} = [\mathbf{t}_1^T \ldots \mathbf{t}_K^T]^T$.

See Res 5.1 and 5.2 of Monahan (2008).

Multivariate Normal distribution

- The pdf of the Normal($\boldsymbol{\mu}, \boldsymbol{\Sigma}$) distribution is given by

$$f(\mathbf{y}; \boldsymbol{\mu}, \Sigma) = (2\pi)^{-p/2}|\boldsymbol{\Sigma}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})\right] \quad \text{for } \mathbf{y} \in \mathbb{R}^p.$$

- The moment generating function is $M_{\mathbf{y}}(\mathbf{t}) = e^{\mathbf{t}^T\boldsymbol{\mu}+\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}/2}$ for all $\mathbf{t} \in \mathbb{R}^p$.
- The pdf of the Normal($\mathbf{0}, \mathbf{I}_p$) distribution is

$$\phi(\mathbf{z}) = (2\pi)^{-p/2}e^{-\mathbf{z}^T\mathbf{z}/2} \quad \text{for } \mathbf{z} \in \mathbb{R}^p.$$

- The mgf of $\mathbf{z} \sim$ Normal($\mathbf{0}, \mathbf{I}_p$) is $M_{\mathbf{z}}(t) = e^{\mathbf{t}^T\mathbf{t}/2}$ for all $\mathbf{t} \in \mathbb{R}^p$.

Exercise: Let $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu}, \mathbf{V})$ and $\mathbf{w} = \mathbf{a} + \mathbf{B}\mathbf{y}$. Show that

$$\mathbf{w} \sim \text{Normal}(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\mathbf{V}\mathbf{B}^T).$$

Result (Distributions of "blocks" of a Normal random vector)

Let $\mathbf{y} \sim \textit{Normal}(\boldsymbol{\mu}, \mathbf{V})$ *and partition* $\mathbf{y}$, $\boldsymbol{\mu}$, *and* $\mathbf{V}$ *as*

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_K \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_K \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \dots & \mathbf{V}_{1K} \\ \vdots & \ddots & \vdots \\ \mathbf{V}_{K1} & \dots & \mathbf{V}_{KK} \end{bmatrix}.$$

Then:

1. $\mathbf{y}_j \sim \textit{Normal}(\boldsymbol{\mu}_j, \mathbf{V}_{jj})$ *for each* j, $j = 1, \dots, K$.
2. $\mathbf{y}_1, \dots, \mathbf{y}_K$ *are mutually independent iff* $\mathbf{V}_{ij} = \mathbf{0}$ *for all* $i \neq j$.

See Cor 5.1 and Res 5.4 of Monahan (2008).

Prove the result.

Exercise: Let $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu}, \mathbf{V})$ and set

$$\mathbf{w}_1 = \mathbf{a}_1 + \mathbf{B}_1\mathbf{y} \quad \text{and} \quad \mathbf{w}_2 = \mathbf{a}_2 + \mathbf{B}_2\mathbf{y}.$$

Show that $\mathbf{w}_1$ and $\mathbf{w}_2$ are independent if and only if $\mathbf{B}_1\mathbf{V}\mathbf{B}_2^T = \mathbf{0}$.

Exercise: Let $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$, $\mathbf{e} \sim \text{Normal}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$.

1. Find the joint distribution of $\hat{\mathbf{y}} = \mathbf{P}_{\mathbf{X}}$ and $\hat{\mathbf{e}} = (\mathbf{I}_n - \mathbf{P}_{\mathbf{X}})\mathbf{y}$.
2. Check whether $\hat{\mathbf{y}}$ and $\hat{\mathbf{e}}$ are independent.

Chi-squared distributions:

- The pdf of the χ^2_ν distribution is given by

$$f_W(w;\nu) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}} w^{\nu/2-1} \exp\left[-\frac{w}{2}\right], \quad \text{for } x > 0.$$

- ν is called the *degrees of freedom*
- mgf: $M_W(t) = (1-2t)^{-\nu/2}$ for $t < 1/2$.
- Let $\chi^2_{\nu,\xi}$ satisfy $P(W > \chi^2_{\nu,\xi}) = \xi$, where $W \sim \chi^2_\nu$.

Let $Z_1, \dots, Z_\nu \overset{\text{ind}}{\sim} \text{Normal}(0,1)$, then

$$W = Z_1^2 + \cdots + Z_\nu^2 \sim \chi^2_\nu.$$

Exercise: Prove the above using mgfs.

t distributions:

- The pdf of the t_ν distribution is given by

$$f_T(t;\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\frac{1}{\sqrt{\nu\pi}}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2} \quad \text{for } t \in \mathbb{R}.$$

- ν is called the *degrees of freedom*
- mgf: does not exist!
- Let $t_{\nu,\xi}$ satisfy $P(T > t_{\nu,\xi}) = \xi$, where $T \sim t_\nu$.

Let Z and W be independent rvs such that $Z \sim \text{Normal}(0,1)$ and $W \sim \chi^2_\nu$, then

$$T = \frac{Z}{\sqrt{W/\nu}} \sim t_\nu.$$

Exercise: Prove above via finding the joint density of (T, U), where $U = W$.

F distributions:

- The pdf of the F_{ν_1,ν_2} distribution is given by

$$f_R(r;\nu_1,\nu_2) = \frac{\Gamma(\frac{\nu_1+\nu_2}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} r^{(\nu_1-2)/2}\left(1+\frac{\nu_1}{\nu_2}r\right)^{-(\nu_1+\nu_2)/2}$$

  for $r > 0$.
- ν_1 and ν_2 are called the *numerator and denominator degrees of freedom*.
- mgf: does not exist!
- Let $F_{\nu_1,\nu_2,\xi}$ satisfy $P(R > F_{\nu_1,\nu_2,\xi}) = \xi$, where $R \sim F_{\nu_1,\nu_2,\xi}$.

Let W_1 and W_2 be independent rvs such that $W_1 \sim \chi^2_{\nu_1}$ and $W_2 \sim \chi^2_{\nu_2}$, then

$$R = \frac{W_1/\nu_1}{W_2/\nu_2} \sim F_{\nu_1,\nu_2}.$$

Exercise: Prove above via finding the joint density of (R, U), where $U = W_2$.

Non-central t distributions:

- The pdf of the $t_{\nu,\phi}$ distribution is given by

$$f_T(t;\nu,\phi) = \frac{e^{-\phi^2/2}}{\sqrt{\pi\nu}\Gamma(\nu/2)} \sum_{k=0}^{\infty} \frac{(2/\nu)^{k/2}(\phi t)^k}{k!} \frac{\Gamma((\nu+k+1)/2)}{(1+(t^2/\nu))^{(\nu+k+1)/2}}, \quad t \in \mathbb{R}.$$

- ν is called the *degrees of freedom*
- ϕ is called the *non-centrality parameter*
- mgf: does not exist!

For $Z \sim \text{Normal}(0,1)$ and $W \sim \chi^2_\nu$ with $Z \perp\!\!\!\perp W$, for any $\phi \in \mathbb{R}$, we have

$$T = \frac{Z+\phi}{\sqrt{W/\nu}} \sim t_{\nu,\phi}.$$

Exercise: For $X_1,\dots,X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu,\sigma^2)$, show that

$$\sqrt{n}(\bar{X}_n - \mu_0)/S_n \sim t_{n-1,\phi}, \qquad \text{with } \phi = \sqrt{n}(\mu-\mu_0)/\sigma.$$

Non-central chi-squared distributions:

- The pdf of the $\chi^2_\nu(\phi)$ distribution is given by

$$f_W(w;\nu,\phi) = \sum_{k=0}^{\infty} \frac{e^{-\phi/2}(\phi/2)^k}{k!} \frac{1}{\Gamma(\frac{\nu+2k}{2})2^{\frac{\nu+2k}{2}}} w^{\frac{\nu+2k}{2}-1} e^{-w/2}, \qquad \text{for } w > 0.$$

- ν is called the *degrees of freedom*.
- ϕ is called the *non-centrality parameter* (ncp).
- mgf: $M_W(t) = \exp(\phi t/(1-2t))(1-2t)^{-\nu/2}$ for $t < 1/2$.

Let $Z_1, \dots, Z_\nu \overset{\text{ind}}{\sim} \text{Normal}(0,1)$, then

$$W = (Z_1 + \mu_1)^2 + \cdots + (Z_\nu + \mu_\nu)^2 \sim \chi^2_\nu(\phi = \mu_1^2 + \cdots + \mu_\nu^2).$$

Note that Monahan (2008) eq. (5.4) parameterizes the ncp differently.

Exercise: Prove the above using mgfs.

Non-central F distributions:

- The pdf of the $F_{\nu_1,\nu_2}(\phi)$ distribution is given by

$$f_R(r;\nu_1,\nu_2) = \sum_{k=0}^{\infty} \frac{e^{-\phi/2}(\phi/2)^2}{k!} \frac{\Gamma(\frac{\nu_1+\nu_2+2k}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_1+2k}{2})} \frac{\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1+2k}{2}} r^{\frac{\nu_1+2k}{2}-1}}{\left(1+r\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1+\nu_2+2k}{2}}}, \quad r > 0.$$

- ν_1 and ν_2 are called the *numerator and denominator degrees of freedom*.
- ϕ is the *non-centrality parameter*.
- mgf: does not exist!

Let W_1 and W_2 be independent rvs such that $W_1 \sim \chi^2_{\nu_1}(\phi)$ and $W_2 \sim \chi^2_{\nu_2}$, then

$$R = \frac{W_1/\nu_1}{W_2/\nu_2} \sim F_{\nu_1,\nu_2}(\phi).$$

Exercise: Prove above via finding the joint density of (R, U), where $U = W_2$.

We will need the following result.

Result (Spectral decomposition for an idempotent matrix)

Let $\mathbf{A}$ be a symmetric $p \times p$ matrix. Then $\mathbf{A}$ is idempotent with rank s if and only if there exists a $p \times s$ matrix $\mathbf{G}$ such that $\mathbf{G}^T\mathbf{G} = \mathbf{I}_s$ and $\mathbf{G}\mathbf{G}^T = \mathbf{A}$.

Prove the result.

Result (Chi-square results for quadratic forms in indep. Normals.)

Let $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_p)$.

1. *We have* $\mathbf{y}^T \mathbf{y}/\sigma^2 \sim \chi^2_p(\phi = \boldsymbol{\mu}^T \boldsymbol{\mu}/\sigma^2)$.
2. *If* $\mathbf{A}$ *symm. idem. w/ rank s, then* $\mathbf{y}^T \mathbf{A} \mathbf{y}/\sigma^2 \sim \chi^2_s(\phi = \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}/\sigma^2)$

Cf. Res 5.9 and 5.14 Monahan (2008).

Prove the results.

Exercise: Let $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$, where $\mathbf{e} \sim \text{Normal}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ and $\mathbf{X}$ has rank r. Define

$$F = \frac{\|\hat{\mathbf{y}}\|^2 / r}{\|\hat{\mathbf{e}}\|^2 / (n - r)},$$

where $\hat{\mathbf{y}} = \mathbf{P}_{\mathbf{X}}\mathbf{y}$ and $\hat{\mathbf{e}} = (\mathbf{I}_n - \mathbf{P}_{\mathbf{X}})\mathbf{y}$. Find the distribution of F.

Result (Chi-square results for quadratic forms in Normals.)

Let $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu}, \sigma^2\mathbf{V})$, $\mathbf{V}$ *a* $p \times p$ *positive definite matrix.*

1. *We have* $\mathbf{y}^T\mathbf{V}^{-1}\mathbf{y}/\sigma^2 \sim \chi^2_p(\phi = \boldsymbol{\mu}^T\mathbf{V}^{-1}\boldsymbol{\mu}/\sigma^2)$
2. *If* $\mathbf{A}$ *symm. and* $\mathbf{AV}$ *is idem. w/rank* s*, then* $\mathbf{y}^T\mathbf{A}\mathbf{y}/\sigma^2 \sim \chi^2_s(\phi = \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}/\sigma^2)$.

Cf. Res 5.10 and 5.15 of Monahan (2008).

Result (Independence of a linear function and a quadratic form)

Let $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu}, \mathbf{V})$ *and* $\mathbf{A}$ *be symm. with rank* s*. If* $\mathbf{BVA} = \mathbf{0}$ *then* $\mathbf{By}$ *and* $\mathbf{y}^T\mathbf{Ay}$ *are independent.*

Prove the result.

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$ and let

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

Show that S_n^2 and $\bar{X}_n$ are independent.

Theorem (Cochran)

Let $\mathbf{y} \sim \text{Normal}(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$ *and let* $\mathbf{A}_1, \dots, \mathbf{A}_K$ *be symmetric, idempotent matrices with ranks* $s_1, \dots, s_K$ *such that* $\sum_{k=1}^K \mathbf{A}_k = \mathbf{I}_n$*. Then*

$$\frac{\mathbf{y}^T \mathbf{A}_1 \mathbf{y}}{\sigma^2}, \dots, \frac{\mathbf{y}^T \mathbf{A}_K \mathbf{y}}{\sigma^2} \quad \textit{are independent}$$

and $\dfrac{\mathbf{y}^T \mathbf{A}_k \mathbf{y}}{\sigma^2} \sim \chi^2_{s_k} \left(\phi = \dfrac{\boldsymbol{\mu}^T \mathbf{A}_k \boldsymbol{\mu}}{\sigma^2} \right)$ *for each* $k = 1, \dots, K$*.*

Prove the result.

Result (Sequential sums of squares by Cochran's theorem)

Let $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$, *where* $\mathbf{e} \sim \text{Normal}(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$ *and* $\mathbf{X} = [\mathbf{X}_0 \ \mathbf{X}_1 \ \ldots \ \mathbf{X}_K]$. *Define*

$$\mathbf{P}_0 = \mathbf{P}_{\mathbf{X}_0}, \quad \mathbf{P}_1 = \mathbf{P}_{[\mathbf{X}_0 \ \mathbf{X}_1]}, \quad \mathbf{P}_2 = \mathbf{P}_{[\mathbf{X}_0 \ \mathbf{X}_1 \ \mathbf{X}_2]}, \quad \ldots \quad \mathbf{P}_K = \mathbf{P}_{\mathbf{X}}$$

and set $s_0 = \operatorname{rank} \mathbf{P}_0$, $s_k = \operatorname{rank}(\mathbf{P}_k - \mathbf{P}_{k-1})$, $k = 1, \ldots, K$. *Then we have*

$$\frac{1}{\sigma^2}\mathbf{y}^T \mathbf{P}_0 \mathbf{y} \sim \chi^2_{s_0}\left(\phi = \frac{1}{\sigma^2}(\mathbf{X}\mathbf{b})^T \mathbf{P}_0 \mathbf{X}\mathbf{b}\right)$$

$$\frac{1}{\sigma^2}\mathbf{y}^T (\mathbf{P}_k - \mathbf{P}_{k-1}) \mathbf{y} \sim \chi^2_{s_k}\left(\phi = \frac{1}{\sigma^2}(\mathbf{X}\mathbf{b})^T (\mathbf{P}_k - \mathbf{P}_{k-1}) \mathbf{X}\mathbf{b}\right), \quad k = 1, \ldots, K$$

$$\frac{1}{\sigma^2}\mathbf{y}^T (\mathbf{I} - \mathbf{P}_K) \mathbf{y} \sim \chi^2_{n - \sum_{k=1}^K s_k},$$

where the above quantities are mutually independent.

Show that this is an application of Cochran's theorem.

We can summarize the sequential sums of squares and their distribution in a table:

Source	SS/σ^2	df	ϕ
$\mathbf{X}_0$	$\frac{1}{\sigma^2}\mathbf{y}^T\mathbf{P}_0\mathbf{y}$	s_0	$\frac{1}{\sigma^2}(\mathbf{Xb})^T\mathbf{P}_0(\mathbf{Xb})$
$\mathbf{X}_k$	$\frac{1}{\sigma^2}\mathbf{y}^T(\mathbf{P}_k - \mathbf{P}_{k-1})\mathbf{y}$	s_k	$\frac{1}{\sigma^2}(\mathbf{Xb})^T(\mathbf{P}_k - \mathbf{P}_{k-1})(\mathbf{Xb})$
$\mathbf{e}$	$\frac{1}{\sigma^2}\mathbf{y}^T(\mathbf{I} - \mathbf{P}_K)\mathbf{y}$	$n - \sum_{k=0}^K s_k$	0

Exercise: Let $Y_i = \mu + \alpha_i + \varepsilon_{ij}$, $\varepsilon_{ij} \overset{\text{ind}}{\sim} \text{Normal}(0, \sigma^2)$, $i = 1, \dots, a$, $j = 1, \dots, n_i$.

Derive the following table summarizing the sequential sums of squares:

Source	SS/σ^2	df	ϕ
Mean	$\dfrac{1}{\sigma^2} n_{.}(\bar{y}_{..})^2$	1	$\dfrac{1}{\sigma^2} n_{.} \left(\mu + n_{.}^{-1} \displaystyle\sum_{i=1}^{a} n_i \alpha_i \right)^2$
Treatment	$\dfrac{1}{\sigma^2} \displaystyle\sum_{i=1}^{a} n_i (\bar{y}_i - \bar{y}_{..})^2$	$a - 1$	$\dfrac{1}{\sigma^2} \displaystyle\sum_{i=1}^{a} n_i \left(\alpha_i - n_{.}^{-1} \sum_{j=1}^{a} n_j \alpha_j \right)^2$
Error	$\dfrac{1}{\sigma^2} \displaystyle\sum_{i=1}^{a} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$	$n_{.} - a$	0

Monahan, J. F. (2008). *A primer on linear models*. CRC Press.