

Outline

1) χ^2 , t and F distributions

2) Canonical linear model

3) General linear model

"Gaussian-adjacent" distributions

If $Z_1, \dots, Z_d \stackrel{iid}{\sim} N(0, 1)$ then

$$V = \sum Z_i^2 \sim \chi_d^2 = \text{Gamma}(\overset{\text{shape}}{d/2}, \overset{\text{scale}}{2})$$

$$\mathbb{E} V = d, \quad \text{Var}(V) = 2d$$

$$\text{CLT}: \frac{V-d}{\sqrt{2d}} \Rightarrow N(0, 1)$$

$$(\text{informal}) \quad \frac{V}{d} \approx N(1, 2/\sqrt{d}) \rightarrow 1$$

If $Z \sim N(0, \sigma^2 \mathbf{I})$ and $V \sim \sigma^2 \chi_d^2$, $Z \perp V$ then

$$\frac{Z}{\sqrt{V/d}} \sim t_d \Rightarrow N(0, 1) \text{ as } d \rightarrow \infty$$

If $V_1 \sim \sigma^2 \chi_{d_1}^2$ and $V_2 \sim \sigma^2 \chi_{d_2}^2$, $V_1 \perp V_2$ then

$$\frac{V_1/d_1}{V_2/d_2} \sim F_{d_1, d_2} \Rightarrow \frac{1}{d_1} \chi_{d_1}^2 \text{ as } d_2 \rightarrow \infty$$

Note if $T \sim t_d$ then $T^2 \sim F_{1, d}$

Recall: $Z \sim N_d(\mu, \Sigma)$, $A \in \mathbb{R}^{k \times d}$, $b \in \mathbb{R}^k$
 $\Rightarrow AZ + b \sim N_k(A\mu + b, A\Sigma A')$

Change of basis (1-sample t-test)

$$X_1, \dots, X_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2) \Leftrightarrow X \sim N(\mu \mathbf{1}_n, \sigma^2 \mathbf{I}_n)$$

$$\text{Let } Q = \left(\overset{1}{q_1} \overset{1}{q_2} \dots \overset{1}{q_n} \right) = \overset{1}{n} \left(\overset{1}{q_1} \overset{n-1}{Q_r} \right)$$

$$\text{where } q_1 = \frac{1}{\sqrt{n}} \cdot \mathbf{1}_n,$$

$q_2, \dots, q_n$ complete orthonormal basis

(e.g. via Gram-Schmidt)

New basis:

$$Z = Q'X = \overset{1}{n-1} \begin{pmatrix} \overset{1}{z_1'} X \\ \overset{1}{Q_r'} X \end{pmatrix} = \begin{pmatrix} \overset{1}{\sqrt{n}} \bar{X} \\ \overset{1}{Q_r' X} \end{pmatrix}$$

$$\begin{aligned} \|Q_r' X\|^2 &= \|Q' X\|^2 - \|q_1' X\|^2 \\ &= \|X\|^2 - n \bar{X}^2 \quad (Q'Q = \mathbf{I}_n) \\ &= (n-1) S^2 \end{aligned}$$

$$Q'X \sim N_n \left(\begin{pmatrix} \sqrt{n} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \sigma^2 \mathbf{I}_n \right)$$

$$Z_1 \sim N(\sqrt{n} \mu, \sigma^2)$$

$$Z_r = Q_r' X \sim N(0, \sigma^2 \mathbf{I}_{n-1})$$

$$\Rightarrow S^2 = \frac{1}{n-1} \|Z_r\|^2 \sim \frac{\sigma^2}{n-1} \chi_{n-1}^2$$

and $S^2 \perp Z_1$ (we already knew, from Basu)

Geometric interp.

$$\begin{aligned} T^2 &= \frac{n \bar{X}^2}{S^2} = \frac{\| \text{Proj}_{Z_n} X \|^2}{\frac{1}{n-1} \| \text{Proj}_{Z_n}^\perp X \|^2} \sim F_{1, n-1} \\ &= \frac{(\text{magnitude of } X \text{ in special dir.})^2}{(\text{average magn. of } X \text{ in resid. dir.s})^2} \end{aligned}$$

Independent of total magnitude (under H_0):

$$n \bar{X}^2 \stackrel{H_0}{\sim} \sigma^2 \chi_1^2 = \text{Gamma}\left(\frac{1}{2}, 2\sigma^2\right)$$

$$(n-1)S^2 \sim \sigma^2 \chi_{n-1}^2 = \text{Gamma}\left(\frac{n-1}{2}, 2\sigma^2\right)$$

$$\|X\|^2 = n \bar{X}^2 + (n-1)S^2 \stackrel{H_0}{\sim} \sigma^2 \chi_n^2 = \text{Gamma}\left(\frac{n}{2}, 2\sigma^2\right)$$

$$\Rightarrow \frac{n \bar{X}^2}{\|X\|^2} \sim \text{Beta}\left(\frac{1}{2}, \frac{n-1}{2}\right), \text{ indep. of } \|X\|^2$$

$$\begin{aligned} & \text{''} \\ & \frac{n \bar{X}^2}{n \bar{X}^2 + (n-1)S^2} \end{aligned}$$

F_{d_1, d_2} related to $\text{Beta}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)$: If $U \sim \text{Beta}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)$ Then $\frac{U/d_1}{(1-U)/d_2} \sim F_{d_1, d_2}$

Canonical Linear Model

Assume $\mathbf{z} = \begin{pmatrix} \mathbf{z}_0 \\ \mathbf{z}_1 \\ \mathbf{z}_r \end{pmatrix} \sim N_n \left(\begin{pmatrix} \mu_0 \\ \mu_1 \\ 0 \end{pmatrix}, \sigma^2 \mathbf{I}_n \right)$

$d_0 = d - d_1$
 $d_r = n - d$

$$\mu_0 \in \mathbb{R}^{d_0}, \mu_1 \in \mathbb{R}^{d_1}, \sigma^2 > 0$$

Test $H_0: \mu_1 = 0$ vs. $H_1: \mu_1 \neq 0$
(or possibly one-sided, if $d_1 = 1$).

Exp. Fam.:

$$p(\mathbf{z}) = e^{\frac{\mu_1'}{\sigma^2} \mathbf{z}_1 + \frac{\mu_0'}{\sigma^2} \mathbf{z}_0 - \frac{1}{2\sigma^2} \|\mathbf{z}\|^2}$$

σ^2 known, $d_1 = 1$:

"Cond. on $\mathbf{z}_0$ ", reject for large(/small/extreme) $\mathbf{z}_1$

$\mathbf{z}_0 \perp \mathbf{z}_1$, test stat is $\mathbf{z}_1 \sim N(\mu_1, \sigma^2)$

$$\frac{\mathbf{z}_1}{\sigma} \stackrel{H_0}{\sim} N(0, 1) \quad (\underline{z\text{-test}})$$

unless we have anisotropic prior on μ_1
 $\downarrow$

σ^2 known, $d_1 \geq 1$: reject for large $\|\mathbf{z}_1\|$

$$\|\mathbf{z}_1\|^2 / \sigma^2 \sim \chi_{d_1}^2 \quad (\underline{\chi^2\text{-test}})$$

σ^2 unknown, $d_1 = 1$:

Cond. on Z_0 , $\|Z\|^2 = \|Z_1\|^2 + \|Z_0\|^2 + \|Z_r\|^2$

Reject for large(/small/extreme) Z_1

$\Leftrightarrow$ Reject for large $Z_1 / \|Z\|$

$\Leftrightarrow$ Reject for large $\frac{Z_1}{\sqrt{\|Z_r\|^2/(n-d)}} \stackrel{H_0}{\sim} t_{d_r}$
(t-test)

σ^2 , $d_1 \geq 1$: Reject for (conditionally) large $\|Z_1\|^2$

$\Leftrightarrow$ Reject for large $\frac{\|Z_1\|^2/d_1}{\|Z_r\|^2/(n-d)} \stackrel{H_0}{\sim} F_{d_1, n-d}$
(F-test)

Here $\|Z_r\|^2/d_r \sim \frac{\sigma^2}{d_r} \chi^2_{n-d}$

functioning as estimator of σ^2

$$E \hat{\sigma}^2 = \sigma^2, \quad \text{Var}(\hat{\sigma}^2) = 2\sigma^2/(n-d)$$

Compare:

$$Z: Z_1/\sigma$$

$$t: Z_1/\hat{\sigma}$$

$$\chi^2: \|Z_1\|^2/\sigma^2$$

$$F: \frac{\|Z_1\|^2/d_1}{\hat{\sigma}^2}$$

Intervals for Canonical Model

How to test $H_0: \mu_1 = \mu_1^0 \in \mathbb{R}^d$?

Problem: μ_1 is not a natural parameter.

Translate problem:

$$\begin{pmatrix} z_0 \\ z_1 - \mu_1^0 \\ z_r \end{pmatrix} \sim N_d \left(\begin{pmatrix} \mu_0 \\ \mu_1 - \mu_1^0 \\ 0 \end{pmatrix}, \sigma^2 I_n \right)$$

Can do same tests with $z_1 - \mu_1^0$ replacing z_1

Invert:

$$\underline{d_1 = 1, \sigma^2 \text{ kn}} \quad \frac{z_1 - \mu_1}{\sigma} \sim N(0, 1) \leadsto \text{CI } z_1 \pm \sigma z_{\alpha/2} = [z_1 - \sigma z_{\alpha/2}, z_1 + \sigma z_{\alpha/2}]$$

$$\underline{d_1 = 1, \hat{\sigma} \text{ unkn}} \quad \frac{z_1 - \mu_1}{\hat{\sigma}} \sim t_{n-d} \leadsto z_1 \pm \hat{\sigma} t_{n-d}^{(\alpha/2)} \quad \left\{ x: \|x\| \leq 1 \right\}$$

$$\underline{d_1 \geq 1, \sigma^2 \text{ kn}}: \left\| \frac{z_1 - \mu_1}{\sigma} \right\| \sim \chi_{d_1}^2 \leadsto z_1 + \sigma \sqrt{c_{\chi^2}(\alpha)} B_1(0)$$

upper- α quantile

$$\underline{d_1 \geq 1, \sigma^2 \text{ unkn}}: \left\| \frac{z_1 - \mu_1}{\hat{\sigma}} \right\| \sim F_{d_1, n-d} \leadsto z_1 + \hat{\sigma} \sqrt{c_F(\alpha)} B_1(0)$$

General Linear Model

Many problems can be put into canonical linear model after change of basis.

Basic setup:

Observe $Y \sim N_n(\theta, \sigma^2 I_n)$, $\sigma^2 > 0$
(known or unknown)

Test $\theta \in \mathcal{H}_0$ vs. $\theta \in \mathcal{H} \setminus \mathcal{H}_0$

where $\mathcal{H}_0 \subseteq \mathcal{H}$ are subspaces of $\mathbb{R}^n$

$\dim(\mathcal{H}_0) = d_0$, $\dim(\mathcal{H}) = d = d_0 + d_1$

Idea: rotate into canonical form

$$Q = \begin{matrix} & \overset{d_0}{\text{orthonormal}} & \overset{d_1}{\text{o.b. for}} & \overset{n-d}{\text{ab. for}} \\ \overset{n}{Q} & [Q_0 & Q_1 & Q_r] \\ & \text{basis for } \mathcal{H}_0 & \mathcal{H} \cap \mathcal{H}_0^\perp & \mathbb{R}^n \cap \mathcal{H}^\perp \end{matrix}$$

$$Z = Q'Y \sim N_n\left(\begin{pmatrix} Q_0'\theta \\ Q_1'\theta \\ 0 \end{pmatrix}, \sigma^2 I_n\right)$$

$$H_0: Q_1'\theta = 0$$

Do z , χ^2 , t , or F -test as appropriate

Ex. Linear Regression $x_i \in \mathbb{R}^d$ fixed

$$y_i = x_i' \beta + \varepsilon_i, \quad \varepsilon_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$$

$$Y \sim N_d(X\beta, \sigma^2 I_n) \quad X = \begin{pmatrix} -x_1' & - \\ \vdots & \\ -x_n' & - \end{pmatrix} \in \mathbb{R}^{d \times n}$$

$$= \begin{pmatrix} | & & | \\ X_1 & \dots & X_d \\ | & & | \end{pmatrix} \quad \text{capital letters}$$

(Assume X has full column rank)

$$\Theta = X\beta \in \mathcal{H} = \text{Span}(X_1, \dots, X_d) \quad (\text{"model space"})$$

$$H_0: \beta_1 = \dots = \beta_{d_1} = 0, \quad (1 \leq d_1 \leq d)$$

$$\Leftrightarrow \Theta \in \mathcal{H}_0 = \text{Span}(X_{d_1+1}, \dots, X_d) \quad (\text{"null model space"})$$
$$(\text{or } \{0\} \text{ if } d_1 = d)$$

Rotate into canonical basis:

Q_0 = orthobasis for $\mathcal{H}_0$

Q_1 = orthobasis for $\mathcal{H} \cap \mathcal{H}_0^\perp$

Q_r = orthobasis for $\mathbb{R}^n \cap \mathcal{H}_0$

$$F\text{-stat} = \frac{\|Q_1' Y\|^2 / d_1}{\|Q_r' Y\|^2 / (n-d)}, \quad t\text{-stat} = \frac{Q_1' Y}{\sqrt{\|Q_r' Y\|^2 / (n-d)}} \quad (\text{if } d_1 = 1)$$

Nice to get more explicit expressions

$$\hat{\beta}_{OLS} = \underset{\beta}{\operatorname{argmin}} \|Y - X\beta\|^2 = (X'X)^{-1}X'Y$$

$$\|z_r\|^2 = \|Q_r'Y\|^2$$

$$= \|\operatorname{Proj}_{\Theta}^{\perp}(Y)\|^2$$

$$= \|Y - \operatorname{Proj}_{\Theta}(Y)\|^2$$

$$= \sum_i (y_i - x_i' \hat{\beta}_{OLS})^2$$

$$r_i = y_i - x_i' \hat{\beta}_{OLS}$$

called residual

$$= \text{Residual sum of squares (RSS)}$$

$n-d$ called residual degrees of freedom

$$\|z_1\|^2 + \|z_r\|^2 = \|[Q, Q_r]'Y\|^2$$

$$= \|\operatorname{Proj}_{\Theta_0}^{\perp}(Y)\|^2$$

$$= \text{RSS}_0 \quad (\text{null RSS})$$

$$F\text{-statistic is } \frac{\|z_1\|^2 / (d-d_0)}{\|z_r\|^2 / (n-d)} = \frac{(\text{RSS}_0 - \text{RSS}) / (d-d_0)}{\text{RSS} / (n-d)}$$

$$d_1 = 1: \quad \text{Let } X_0 = (X_2 \dots X_d) \in \mathbb{R}^{d \times n}$$

$$q_1 = \frac{X_{1\perp}}{\|X_{1\perp}\|},$$

$$\text{where } X_{1\perp} = X_1 - \text{Proj}_{\mathcal{W}_0}(X_1)$$

$$= X_1 - X_0 (X_0' X_0)^{-1} X_0' X_1$$

$$= X_1 - X_0 \gamma$$

Reparametrize:

$$\begin{aligned} \Theta = X\beta &\Leftrightarrow \Theta = X_{1\perp} \beta_1 + X_0 \overbrace{(\beta_{-1} + \gamma \beta_1)}^{\delta} \\ &= \begin{bmatrix} X_{1\perp} & X_0 \end{bmatrix} \begin{pmatrix} \beta_1 \\ \delta \end{pmatrix} \end{aligned}$$

OLS solution in new parametrization:

$$\begin{aligned} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\delta} \end{pmatrix} &= \left(\begin{bmatrix} X_{1\perp} & X_0 \end{bmatrix}' \begin{bmatrix} X_{1\perp} & X_0 \end{bmatrix} \right)^{-1} \begin{bmatrix} X_{1\perp} & X_0 \end{bmatrix}' Y \\ &= \begin{bmatrix} X_{1\perp}' X_{1\perp} & 0 \\ 0 & X_0' X_0 \end{bmatrix}^{-1} \begin{pmatrix} X_{1\perp}' Y \\ X_0' Y \end{pmatrix} = \begin{pmatrix} X_{1\perp}' Y / \|X_{1\perp}\|^2 \\ (X_0' X_0)^{-1} X_0' Y \end{pmatrix} \end{aligned}$$

$$\hat{\beta}_1 = X_{1\perp}' Y / \|X_{1\perp}\|^2, \quad \text{s.e.}(\hat{\beta}_1) \triangleq \sqrt{\text{Var}(\hat{\beta}_1)} = \sigma / \|X_{1\perp}\|$$

$$t\text{-statistic: } \frac{q_1' Y}{\sqrt{RSS/(n-d)}} = \frac{\hat{\beta}_1}{\hat{\sigma} / \|X_{1\perp}\|} = \frac{\hat{\beta}_1}{\text{s.e.}(\hat{\beta}_1)}$$

Ex: Two-sample t-test (equal variance)

$$Y_1, \dots, Y_m \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2) \quad Y_{m+1}, \dots, Y_{n+m} \stackrel{\text{iid}}{\sim} N(\nu, \sigma^2)$$

$$\text{Model: } \Theta = \mathbb{E}Y = \begin{pmatrix} m \mathbf{1}_m \\ \nu \mathbf{1}_n \end{pmatrix} \Leftrightarrow \Theta \in \text{Span} \left(\begin{pmatrix} \mathbf{1}_m \\ -\mathbf{1}_n \end{pmatrix}, \mathbf{1}_{n+m} \right)$$

$$H_0: \mu_1 = \mu_2 \Leftrightarrow \Theta \in \text{Span}(\mathbf{1}_{n+m})$$

$$d_0 = 1, \quad d = 2, \quad d_r = n+m-2$$

$$\text{Orthogonalize } \begin{pmatrix} \mathbf{1}_m \\ -\mathbf{1}_n \end{pmatrix} \rightsquigarrow \left\{ \begin{matrix} m \\ n \end{matrix} \right\} \begin{pmatrix} 1/m \\ \vdots \\ 1/m \\ -1/n \\ \vdots \\ -1/n \end{pmatrix}$$

$\Rightarrow$ Reject for large

$$\frac{\frac{1}{m} \sum_{i \in m} Y_i - \frac{1}{n} \sum_{i \geq m} Y_i}{\sqrt{\frac{1}{m} + \frac{1}{n}} \cdot \sqrt{\text{RSS}/(n+m-2)}} = \frac{\bar{Y}_1 - \bar{Y}_2}{\hat{\sigma} \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

Ex. One-way ANOVA: (fixed effects)

$$Y_{k,i} \stackrel{\text{ind.}}{\sim} \mu_k + \varepsilon_{k,i} \quad \varepsilon_{k,i} \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$$

$$k = 1, \dots, m \quad i = 1, \dots, n$$

$$H_0: \mu_1 = \dots = \mu_m = \mu$$

$$\bar{Y}_k = \frac{1}{n} \sum_i Y_{k,i} \quad S_k^2 = \frac{1}{n-1} \sum_i (Y_{k,i} - \bar{Y}_k)^2$$

$$\bar{Y} = \frac{1}{mn} \sum_k \sum_i Y_{k,i} \quad S_0^2 = \frac{1}{mn-1} \sum_k \sum_i (Y_{k,i} - \bar{Y})^2$$

$$d_0 = 1, \quad d = m, \quad d_r = m(n-1)$$

$$RSS = \sum_{k,i} (Y_{k,i} - \bar{Y}_k)^2 = \|Y\|^2 - n \sum_k \bar{Y}_k^2$$

$$RSS_0 = \sum_{k,i} (Y_{k,i} - \bar{Y})^2 = \|Y\|^2 - mn \bar{Y}^2$$

$$\begin{aligned} RSS_0 - RSS &= n \left(\sum_k \bar{Y}_k^2 - m \bar{Y}^2 \right) \\ &= n \sum_k (\bar{Y}_k - \bar{Y})^2 \end{aligned}$$

$$F\text{-stat} = \frac{\frac{n}{m-1} \sum_k (\bar{Y}_k - \bar{Y})^2 \leftarrow \text{"between" variance}}{\frac{1}{m(n-1)} \sum_k \sum_i (Y_{k,i} - \bar{Y}_k)^2 \leftarrow \text{"within" variance}}$$