Rep. multivariate normal and distⁿ of quadratic forms.

Random vector $\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$ e.g. $\vec{X}_i$ represents measurements of person i; height, weight, blood pressure, etc.

Univar: $E(X_1) = \mu_1 \quad E(X_2) = \mu_2$ etc.

$Var(X_1) = E((X_1-\mu_1)^2) = \sigma_1^2 \quad Var(X_2) = \sigma_2^2$ etc.

$Cov(X_1, X_2) = E((X_1-\mu_1)(X_2-\mu_2)) = \sigma_{12}$

Multivar: $E(\vec{X}) = \vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$

$$Cov(\vec{X}) = E((\vec{X}-\vec{\mu})(\vec{X}-\vec{\mu})^T) = \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & & \vdots \\ & & \ddots & \\ & & & \sigma_q^2 \end{bmatrix}$$

A $q \times p$ const. matrix, $\vec{c}$ constant q-vector

$\vec{Y} = A\vec{X} + \vec{c} \qquad E(\vec{Y}) = A\vec{\mu} + \vec{c} \qquad Cov(\vec{Y}) = A\Sigma A^T$

Useful transf.: $\vec{Z} = \Sigma^{-1/2}(\vec{X} - \vec{\mu})$ (recall $\frac{X_1 - \mu_1}{\sqrt{\sigma^2}} \rightarrow$ mean 0 var 1)

$E(\vec{Z}) = \vec{0} \quad Cov(\vec{Z}) = I$

nb: $\underset{\text{pos. def.}}{\Sigma} = \overset{\text{ort.}}{P} \underset{\text{diag.}}{\Lambda} P^T$

$\Sigma^{-1/2} = P\Lambda^{-1/2}P$

Multivariate normal

$\vec{X} \sim MVN(\vec{\mu}, \Sigma)$

- then $X_1 \sim N(\mu_1, \sigma_1^2)$ etc
- If $\text{cov}(X_1, X_2) = 0$ then X_1, X_2 independent
- $\vec{a}$ const. p-vector,
  then $\vec{a}^T \vec{X}$ univar. normal
- $A, \vec{c}$ const., $A\vec{X} + \vec{c}$ multivar normal
  $\Sigma^{-1/2}(\vec{X} - \vec{\mu}) \sim MVN(\vec{0}, I)$
- Independent lin. trans
  $A\vec{X}$ and $B\vec{X}$ independent if $A \Sigma B^T = 0$
  $\text{cov}(A\vec{X}, B\vec{X})$
- cond. distn.
  $\vec{X} = \begin{bmatrix} \vec{X}_1 \\ \vec{X}_2 \end{bmatrix}$ $\vec{\mu} = \begin{bmatrix} \vec{\mu}_1 \\ \vec{\mu}_2 \end{bmatrix}$ $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$

$$\vec{X}_1 | \vec{X}_2 = \vec{x}_2 \sim MVN\left(\vec{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\vec{x}_2 - \vec{\mu}_2),\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$$

Exam 2016

$\vec{Y} = \underset{\substack{\uparrow \\ 2\times 2}}{A}\vec{X}$, bivariate normal

$$E(\vec{Y}) = AE(\vec{X}) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

$$\text{cov}(\vec{Y}) = A\Sigma A^T = \underbrace{\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}}_{\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

$A\vec{X}, B\vec{X}$ ind if $A\Sigma B^T = 0$

$$A\Sigma B^T = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & b \\ a & 1 \end{pmatrix} = \begin{pmatrix} 2-a & b-1 \\ -2+a & -b+1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$a=2 \quad b=1$

Use in linear regression

$$\vec{Y} = X\vec{\beta} + \vec{\varepsilon}$$

NB: X is here a matrix of <u>constants</u> $(n \times p)$

$\vec{\beta}$ is a p-vector of (unknown) constants.

$\vec{\varepsilon} \sim MVN_n(\vec{0}, \underbrace{\sigma^2 I}_{\Sigma})$ ← all covar 0 → independence

$$\vec{Y} \sim MVN_n(X\vec{\beta}, \sigma^2 I)$$

$\rightarrow \quad Y_1 \sim N(\vec{X}_1^T\vec{\beta}, \sigma^2)$ etc

$Y_1 | Y_2 = y_2 \sim N(\vec{X}_1^T\vec{\beta}, \sigma^2)$ (ind.)

Estimator: $\hat{\vec{\beta}} = (X^TX)^{-1}X^T\vec{Y}$

$$E(\hat{\vec{\beta}}) = (X^TX)^{-1}X^TX\vec{\beta} = \vec{\beta}$$

$$\mathrm{cov}(\hat{\vec{\beta}}) = (X^TX)^{-1}X^T \sigma^2 I X (X^TX)^{-1}.$$
$$= \sigma^2(X^TX)^{-1}$$

$$\hat{\vec{\beta}} \sim MVN(\vec{\beta}, \sigma^2(X^TX)^{-1})$$

residuals $\hat{\varepsilon} = \vec{Y} - \hat{\vec{Y}}$
$= (I-H)\vec{Y}$

$\hat{Y} = X\hat{\beta} = \underbrace{X(X^TX)^{-1}X^T}_{H}\vec{Y}$

"puts the hat on Y!"

H, $I-H$, symm. idempotent "projection matrices"

$E(\hat{\varepsilon}) = (I-H)X\vec{\beta} = X\vec{\beta} - X\vec{\beta} = \vec{0}$ $\quad HX = X$

$\mathrm{Cov}(\hat{\varepsilon}) = (I-H)\sigma^2 I(I-H)^T$ ← symm

$= \sigma^2(I-H)(I-H) = \sigma^2(I-H)$ idemp.

$(I-H)$ $n \times n$ but $\mathrm{rank}(I-H) = n-p$, $(I-H)$ not invertible.

$$\hat{\varepsilon} \sim MVN(\vec{0}, \sigma^2(I-H))$$

$\hat{\varepsilon}$ and $\hat{\beta}$ independent

Use $A\vec{X}, B\vec{X}$ ind if $A\Sigma B^T = 0$

here $A = (X^T X)^{-1} X^T$

$B = (I - H)$

$\Sigma = \sigma^2 I$

$$A\Sigma B^T = \sigma^2 (X^T X)^{-1} X^T (I-H)$$
$$= \sigma^2 (X^T X)^{-1} (X^T - X^T)$$
$$= 0$$

$HX = X$

$\hat{\sigma}^2 = \hat{\varepsilon}^T \hat{\varepsilon} / n-p$, ind. of $\hat{\beta}$.

$\uparrow$

quadratic form.

Dist^n of quadratic forms

$$\vec{Y} \sim MVN_n(\vec{\mu}, \sigma^2 I) \qquad \text{e.g. } \vec{\mu} = X\vec{\beta}$$

If A symm., idem. $n \times n$ mat. rank r,

then

$$\frac{(\vec{Y}-\vec{\mu})^T A (\vec{Y}-\vec{\mu})}{\sigma^2} \sim \chi^2_r$$

(because we end up summing over r squared standard normal)

If B as A, rank s, $AB = 0$, then

$$\frac{\dfrac{(\vec{Y}-\vec{\mu})^T A (\vec{Y}-\vec{\mu})}{\cancel{\sigma^2}} \Big/ r}{\dfrac{(\vec{Y}-\vec{\mu}) B (\vec{Y}-\vec{\mu})}{\cancel{\sigma^2}} \Big/ s} \sim F_{r,s}$$

Exam 2024

$\vec{Y} \sim MVN_n(X\vec{\beta}, I)$

$X(X^TX)^{+}X^T = H$ symm, idemp. rank p (= rank X)

$QF \sim \chi^2_p$ $E(QF) = p$ $Var(QF) = 2p$

Use of QF in lin. reg.

$SST = SSR + SSE$ ← $\hat{\varepsilon}^T\hat{\varepsilon}$

$\vec{Y}^T(I - \frac{1}{n}J)\vec{Y} = \vec{Y}^T(H - \frac{1}{n}J)\vec{Y} + \vec{Y}(I-H)\vec{Y}$

$J = \vec{1}\vec{1}^T$

ranks: $n-1 = p-1 + n-p$

F-test for sign. of regression

$H_0: \beta_1 = \beta_2 = \cdots = \beta_p = 0$ (not intercept)

Dist^n of SSR/σ^2 under H_0:

$(\vec{Y} - X\vec{\beta})^T(H - \frac{1}{n}J)(\vec{Y} - X\vec{\beta})/\sigma^2$

$(\vec{Y} - \beta_0\vec{1})^T(H - \frac{1}{n}J)(\vec{Y} - \beta_0\vec{1})/\sigma^2 =$

$\vec{Y}^T(H - \frac{1}{n}J)\vec{Y}/\sigma^2 \sim \chi^2_{p-1}$

$\frac{1}{n}J\vec{1} = \vec{1}$

$H\vec{1} = \vec{1}$

$\rightarrow (H - \frac{1}{n}J)\vec{1} = 0$

Dist^n of SSE/σ^2:

use $(I-H)X = 0$

$(\vec{Y} - X\vec{\beta})^T(I-H)(\vec{Y} - X\vec{\beta})/\sigma^2 = \vec{Y}^T(I-H)\vec{Y}/\sigma^2 \sim \chi^2_{n-p}$

$$F = \frac{SSR/p-1}{SSE/n-p} \sim F_{p-1, n-p}$$

using $(H - \frac{1}{n}J)(I-H) = 0$

H_0 (otherwise SSR behaves differently)

Why 'small' values of F under H_0, large values of F under H_a?