样本协方差: $Cov(x, y)=\sum_{i=1}^n(x_i-\bar{x})(y_i-\bar{y})/(n-1)$

大样本定理.: $Cov(x, y)=E((X-E(X))(Y-E(Y)))(n\to\infty)$

相关系数: $r = Cov(x, y)/\sqrt{Var(x)Var(y)}=^{\Delta}S_{xy}/\sqrt{S_{xx}S_{yy}}$

In regression, one variable is considered independent variable(predictor); other dependent variable(outcome).

线性回归假设: I. The relationship between x and y is linear; II. y is distributed normally at each value x; III. The variance of y at every value of x is the same (homogeneity); IV. The observation is independent.

简单线性回归模型

假设: $E(\varepsilon_i)=0$; $Var(\varepsilon_i)=\sigma^2$; $Cov(\varepsilon_i, \varepsilon_j)=0$

最小二乘法: $\min\sum_{i=1}^n(y_i-y_i^{hat})^2$

$\beta_1^{hat}=\sum_{i=1}^n(x_i-\bar{x})(y_i-\bar{y})/\sum_{i=1}^n(x_i-\bar{x})^2=^{\Delta}S_{xy}/S_{xx}$

$\beta_0^{hat}=\bar{y}-\beta_1^{hat}\bar{x}$

$S^2=SSE/(n-2)=\sum_{i=1}^n(y_i-y^{hat})^2/(n-2)$ Unbiased for σ^2

极大似然法: $S_{MLE}^2=SSE/n=(n-2)S^2/n$

$\beta_1^{hat}\sim N(\beta_1, \sigma^2/S_{xx})$, $\beta_0^{hat}\sim N(\beta_0, \sigma^2(1/n+\bar{x}^2/\sum_{i=1}^n(x_i-\bar{x})^2))$

$(n-2)S_{LSE}^2/\sigma^2\sim\chi^2(n-2)$.考虑到$\beta_0^{hat}$ 与 S^2 独立，所以

$(\beta_0^{hat}-\beta_0)/(S\sqrt{1/n+\bar{x}^2/S_{xx}})\sim t(n-2)$

β_0 的$(1-\alpha)$CI: $[\beta_0^{hat}\pm S\, t(\alpha/2, n-2)/\sqrt{1/n+\bar{x}^2/S_{xx}}]$

β_1 双边 t 检验: H_0: $\beta_1=0$ vs. H_1: $\beta_1\neq0$

$t=(\beta_1^{hat}-\beta_1)\sqrt{S_{xx}}/S\sim t(n-2)$, reject H_0 if $|t| > t(\alpha/2, n-2)$

$(1-\alpha)$CI: $[\beta_1^{hat}-S\, t(\alpha/2, n-2)/\sqrt{S_{xx}}, \beta_1^{hat}+S\, t(\alpha/2, n-2)/\sqrt{S_{xx}}]$

Est. of Mean of the Response Varibale for a Given Level of x:

$y_h^{hat}\sim N(\beta_0+\beta_1x_h, \sigma^2[1/n+(x_h-\bar{x})^2/\sum_i(x_i-\bar{x})^2])$ Est. of $E(y_h)$

CI of $E(y_h)$:$[y_h^{hat}\pm t(\alpha/2, n-2)S\sqrt{1/n+(x_h-\bar{x})^2/\sum_i(x_i-\bar{x})^2}]$

Prediction of a new observation

$y_h-y_h^{hat}\sim N(0, \sigma^2[1+1/n+(x_h-\bar{x})^2/\sum_i(x_i-\bar{x})^2])$

Predictive Itv.:$[y_h\pm t_{\alpha/2, n-2}S\sqrt{1+1/n+(x_h-\bar{x})^2/\sum_i(x_i-\bar{x})^2}]$

Analysis of Variance Approach to Regression Analysis

$SST=\sum(y_i-\bar{y})^2=\sum(y_i-y_i^{hat})^2+\sum(y_i^{hat}-\bar{y})^2=SSE+SSR$

SSE Sum of Squares of Residual (Unexplained Variation)

SSR Sum of squares due to Regression (Explained Var.)

Analysis of Variation (ANOVA) Table

	SS	Df	MS	F
Regression	SSR	1	MSR=SSR	MSR/MSE
Error	SSE	n-2	MSE=SSE/(n-2)	
Total	SST	n-1		

F 检验: H_0: $\beta_1=0$(No linear relationship) vs. H_1: $\beta_1\neq0$

Reject H_0 if $F>F(\alpha, 1, n-2)$

Coefficient of Determination (R-squared)

$R^2=SSR/SST = 1-(SSE/SST)$

$r=\pm\sqrt{R^2}$ True for Simple Linear Regression only

满秩分解 Factorization: $A_{p*q}=K_{p*r}L_{r*q}$, $r=r(A)$, K, L full rank

Let $B=A_{22}-A_{21}A_{11}^{-1}A_{12}$; $D=A_{11}-A_{12}A_{22}^{-1}A_{21}$

Remark 3.1 A^{-1}=

$A_{11}^{-1}+A_{11}^{-1}A_{12}B^{-1}A_{21}A_{11}^{-1}$	$-A_{11}^{-1}A_{12}B^{-1}$
$-B^{-1}A_{21}A_{11}^{-1}$	B^{-1}

=

D^{-1}	$-D^{-1}A_{12}A_{22}^{-1}$
$-A_{22}^{-1}A_{21}D^{-1}$	$A_{22}^{-1}+A_{22}^{-1}A_{21}D^{-1}A_{12}A_{22}^{-1}$

$(A+cc')^{-1}=A^{-1}-A^{-1}cc'A^{-1}/(1+c'A^{-1}c)$

幂等矩阵 Idempotent matrix $A^2=A$

对于幂等矩阵 A，有 $r(A) = tr(A)$

证明：考虑 A 的满秩分解 $A=BC$, $A^2=BCBC=BC$, $CB=I_r$

所以 $tr(A)=tr(BC)=tr(CB)=r=rank(A)$.

幂等矩阵的特征值均为 0 或 1.

Theorem 秩 r 矩阵 A_{n*p}假如可划分为$[[A_{11}\ A_{12}][A_{21}\ A_{22}]]$，且 A_{11} 是 r 阶满秩矩阵，则 $A^-=[[A_{11}^{-1}\ 0][0\ 0]]$

类似地，若 $A_{12}/A_{21}/A_{22}$ 分别是 r 阶满秩矩阵，则 A^-分别可以是$[[0\ 0][A_{12}^{-1}\ 0]]/[[0\ A_{21}^{-1}][0\ 0]]/[[0\ 0][0\ A_{22}^{-1}]]$

广义逆 Generalized Inverse A^-: $AA^-A = A$; $rank(A)=k$, A_{m*n}

$rank(A^-)\geqslant rank(A^-A) = rank(AA^-) = k$

A^-A 和 AA^- 是幂等的

$A^-A = I \leftrightarrow rank(A) = n$

$AA^- = I \leftrightarrow rank(A) = m$

$tr(A^-A) = tr(AA^-) = k$

A'^-是 A'的广义逆

$K = A(A'A)^-A'$对任意$(A'A)^-$均不变

$K = K'$; $K = K^2$ (Symmetric, Idempotent)

$rank(K) = rank(A) = tr(K)$

$KA = A$; $A'K = A$

$(X'X)^-X'$ is a g-inverse of X for any g-inverse of X'X

$X(X'X)^-$ is a g-inverse of X' for any g-inverse of X'X

Moore-Penrose Inverse A^+ for A_{m*n}

For every matrix A, there exists A^+ that satisfies

I. AA^+ 对称

II. A^+A 对称

III. $AA^+A = A$

IV. $A^+AA^+ = A^+$

Properties:

$(A')^+ = (A^+)'$; $(A^+)^+ = A$; $rank(A^+) = A$

A symmetric $\leftrightarrow$ A^+ symmetric

A nonsingular $\leftrightarrow A^{-1} = A^+$

A symmetric idempotent $\to A^+ = A$

$rank(A) = m \to A^+ = A'(AA')^{-1}$, $AA^+ = I$

$rank(A) = n \to A^+ = (A'A)^{-1}A'$, $A^+A = I$

A^+A, AA^+, $I-AA^+$, $I-A^+A$ all symmetric idempotent

Vector and Matrix Calculus:

x 是向量: $u = f(\mathbf{x})$

$f(\mathbf{x}) = \mathbf{a}'\mathbf{x} = \mathbf{x}'\mathbf{a} \to \partial u/\partial\mathbf{x} = \mathbf{a}$

$f(\mathbf{x}) = \mathbf{x}'A\mathbf{x} \to \partial u/\partial\mathbf{x} = 2A\mathbf{x}$

X 是矩阵: $u = f(X_{p*p})$, $\partial u/\partial X = [\partial u/\partial x_{ij}]$

$f(X) = tr(XA) \to \partial u/\partial X = A + A' - diag(A)$

$f(X) = \ln|A| \to \partial u/\partial X = 2X^{-1} - diag(X^{-1})$

x 是常数: $A_{n*n} = [a_{ij}]$, a_{ij} 是 x 的函数,

$f(x) = A^{-1} \to \partial u/\partial x = -A^{-1}\,\partial A/\partial x\, A^{-1}$

$f(x) = \ln|A| \to \partial u/\partial x = tr(A^{-1}\partial A/\partial x)$

Random Vector abd Multivariate Normal Distribution

$Cov(\mathbf{Y})=\Sigma=E((\mathbf{Y}-E(\mathbf{Y}))(\mathbf{Y}-E(\mathbf{Y})'))=[\sigma_{ij}]=[Cov(Y_i, Y_j)]\geq\mathbf{0}$

$Cov(A\mathbf{Y})=ACov(\mathbf{Y})A'$ A 是常数矩阵

$Cov(A\mathbf{X}, B\mathbf{Y}) = ACov(\mathbf{X}, \mathbf{Y})B'$

Mahalabolis Distance: $(\mathbf{Y}-\mu)'\Sigma^{-1}(\mathbf{Y}-\mu)$ (Standardized Dist.)

Generalized Variance = $|\Sigma|$

Remark 4.1 一系列观察 $\mathbf{y}_i=[y_{i1} \ldots y_{ip}]\sim$(i.i.d.)$(\mu, \Sigma)$

样本均值向量 $\mu^{hat}=\bar{\mathbf{y}}=[y_1 \ldots y_p]$, $\bar{y}_j=\sum_{i=1}^n y_{ij}/n$

样本协方差矩阵 $\Sigma^{hat}_{p*p}=S=\sum_{i=1}^n(\mathbf{y}_i-\bar{\mathbf{y}})(\mathbf{y}_i-\bar{\mathbf{y}})'$

相关系数矩阵 $\Omega = [\rho_{ij}] = [\sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}]$

Remark 4.2 令 A 为对称矩阵, $E(\mathbf{Y}'A\mathbf{Y})=tr(A\Sigma)+\mu'A\mu$

证明: $E(\mathbf{Y}'A\mathbf{Y})=E((\mathbf{Y}-\mu)'A(\mathbf{Y}-\mu))+\mu'A\mu$

$E((\mathbf{Y}-\mu)'A(\mathbf{Y}-\mu))=E\,tr((\mathbf{Y}-\mu)'A(\mathbf{Y}-\mu))=E\,tr(A(\mathbf{Y}-\mu)(\mathbf{Y}-\mu)')$

$=tr(A\,E((\mathbf{Y}-\mu)(\mathbf{Y}-\mu)'))=tr(A\Sigma)$

Theorem: $\mathbf{Y}_1, \ldots, \mathbf{Y}_m$ 独立$\to g_1(\mathbf{Y}_1), \ldots, g_m(\mathbf{Y}_m)$独立.

MGF: $M_\mathbf{Y}(\mathbf{t}) = E(e^{\mathbf{t}'\mathbf{Y}})$

Let $\mathbf{Y}_{p*1} \sim N(\mu, \Sigma)$, $M_\mathbf{Y}(\mathbf{t})=e^{\mathbf{t}'\mu+\mathbf{t}'\Sigma\mathbf{t}/2}$

$f_\mathbf{Y}(\mathbf{y})=|\Sigma|^{-1/2}(2\pi)^{-p/2}\exp(-(\mathbf{y}-\mu)'\Sigma^{-1}(\mathbf{y}-\mu)/2)$

Remark 4.4 令 B 为常数矩阵，**C** 为常数向量，则

$\mathbf{X}=B\mathbf{Y}+\mathbf{C}\sim N(B\mu+\mathbf{C}, B\Sigma B')$

证明: $M_\mathbf{X}(\mathbf{t})=E(e^{\mathbf{t}'(B\mathbf{Y}+\mathbf{C})})=E(e^{\mathbf{t}'B\mathbf{Y}})e^{\mathbf{t}'\mathbf{C}}$. 记 $\mathbf{t}^*=B'\mathbf{t}$，则

$M_\mathbf{X}(\mathbf{t})=E(e^{\mathbf{t}^{*\prime}\mathbf{Y}})e^{\mathbf{t}'\mathbf{C}}=M_\mathbf{Y}(\mathbf{t}^*)e^{\mathbf{t}'\mathbf{C}}=\exp(\mathbf{t}^{*\prime}\mu+\mathbf{t}^{*\prime}\Sigma\mathbf{t}^*/2+\mathbf{t}'\mathbf{C})$

$=\exp(\mathbf{t}'(B\mu+\mathbf{C})+\mathbf{t}'B\Sigma B'\mathbf{t}/2)$. 即证.

令 $\mathbf{Y}=[[\mathbf{Y}_1]\ [\mathbf{Y}_2]] \sim N([[\mu_1][\mu_2]], [[\Sigma_{11}\ \Sigma_{12}]\ [\Sigma_{21}\ \Sigma_{22}]]$,

$\mathbf{Y}_1 \sim N(\mu_1, \Sigma_{11})$

$\mathbf{Y}_1|(\mathbf{Y}_2 = \mathbf{y}_2) \sim N(\mu_1+\Sigma_{12}\Sigma_{22}^{-1}(\mathbf{y}_2-\mu_2), \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$

$\mathbf{Y}_1$ 和 $\mathbf{Y}_2$ 独立当且仅当$\Sigma_{12}=0$

偏相关 Partial Correlation:

$\mathbf{v}=[[\mathbf{y}][\mathbf{x}]]\sim N_q([[\mu_1][\mu_2]], [[\Sigma_{yy}\ \Sigma_{yx}]\ [\Sigma_{xy}\ \Sigma_{xx}]])$

$\Sigma_{y.x}=\Sigma_{yy}-\Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}=[\sigma_{ij.r\ldots q}]$, $D_{y.x}=[diag(\Sigma_{y.x})]^{1/2}$

$\rho_{ij.r\ldots q}=\sigma_{ij.r\ldots q}/\sqrt{\sigma_{ii.r\ldots q}\sigma_{jj.r\ldots q}}$, $\Omega_{y.x}=D_{y.x}^{-1}\Sigma_{y.x}D_{y.x}^{-1}$

二次型 Quadratic Forms: 令 $\mathbf{x}\sim N(\mu, \Sigma)$，A 是对称矩阵，则

$M_{\mathbf{x}'A\mathbf{x}}(t)=|I-2tA\Sigma|^{-1/2}\exp(-\mu'(I-(I-2tA\Sigma)^{-1})^{-1}\Sigma^{-1}\mu/2)$

$E(\mathbf{x}'A\mathbf{x})=tr(A\Sigma)+\mu'A\mu$, $Var(\mathbf{x}'A\mathbf{x})=2tr[(A\Sigma)^2]+4\mu'A\Sigma A\mu$

Non-central χ^2 distribution

$\mathbf{x}\sim N(\mathbf{0}, I_n)$, $\mathbf{x}'\mathbf{x}\sim\chi^2(n)$

$\mathbf{x}\sim N(\mu, I_n)$, $u=\mathbf{x}'\mathbf{x}\sim\chi^2(n, \lambda)$, $\lambda=\mu'\mu/2$

$f(u)=e^{-\lambda}\sum_{k=0}^{\infty}\lambda^k u^{n/2+k-1}e^{-u/2}/(k!2^{n/2+k}\Gamma(n/2+k))$, $\mu>0$, $\lambda\geq0$

$M_u(t)=(1-2t)^{-n/2}\exp(-\lambda[1-(1-2t)^{-1}])$

$E(u)=n+2\lambda$, $Var(u)=2n+8\lambda$

$u_i\sim\chi^2(n_i, \lambda_i)$ 独立, $\sum u_i\sim\chi^2(\sum n_i, \sum\lambda_i)$

Non-central F distribution

$u_1\sim\chi^2(p_1, \lambda)$, $u_2\sim\chi^2(p_2, 0)$, $w=u_1p_2/u_2p_1\sim F(p_1, p_2, \lambda)$

$E(w)=(1+2\lambda/p_1)p_2/(p_2-2)$

Non-central t distribution

$z\sim N(\mu, 1)$, $u\sim\chi^2(n)$, z is independent of u, then

$t=z/\sqrt{u/n}\sim$non-centered t distribution

Theorem 5.1

$\mathbf{x}_{p*1}\sim N(\mu, \Sigma)$, A 对称, $r=rank(A)$, $\lambda=\mu'A\mu/2$ 则

$q = \mathbf{x}'A\mathbf{x} \sim\chi^2(r, \lambda) \leftrightarrow A\Sigma$幂等

推论

$\mathbf{x}\sim N(\mathbf{0}, I)$，则 $\mathbf{x}'A\mathbf{x}\sim\chi^2(r) \leftrightarrow$ A 幂等

$\mathbf{x}\sim N(\mathbf{0}, \Sigma)$，则 $\mathbf{x}'A\mathbf{x}\sim\chi^2(r) \leftrightarrow A\Sigma$幂等

$\mathbf{x}\sim N(\mu, \sigma^2 I)$, then $\mathbf{x}'\mathbf{x}/\sigma^2\sim\chi^2(n, \mu'\mu/(2\sigma^2))$

$\mathbf{x}\sim N(\mu, I)$，则 $\mathbf{x}'A\mathbf{x}\sim\chi^2(r, \mu'A\mu/2) \leftrightarrow$ A 幂等

Theorem 5.2

$\mathbf{x}\sim N(\mu, \Sigma)$, A 对称，则 $\mathbf{x}'A\mathbf{x}$ 与 $B\mathbf{x}$ 独立 $\leftrightarrow B\Sigma A=O$.

Theorem 5.3

$\mathbf{x}\sim N(\mu, \Sigma)$, A 和 B 对称, $r=rank(A)$，则

$\mathbf{x}'A\mathbf{x}$ 和 $\mathbf{x}'B\mathbf{x}$ 均幂等$\leftrightarrow A\Sigma B=B\Sigma A=O$.

Remark 4.3 证明: $\mathbf{Y}_{p*1} \sim N(\mu, \Sigma)$，则 $M_\mathbf{Y}(\mathbf{t})=e^{\mathbf{t}'\mu+\mathbf{t}'\Sigma\mathbf{t}/2}$.

解: $M_\mathbf{Y}(\mathbf{t})=|\Sigma|^{-1/2}(2\pi)^{-p/2}\int\exp(\mathbf{t}'\mathbf{y}-(\mathbf{y}-\mu)'\Sigma^{-1}(\mathbf{y}-\mu)/2)d\mathbf{y}$

$(*)=(\mathbf{y}-\mu)'\Sigma^{-1}(\mathbf{y}-\mu)-2\mathbf{t}'\mathbf{y}$ (记$\mu^*=\Sigma\mathbf{t}+\mu$)

$=\mathbf{y}'\Sigma^{-1}\mathbf{y}-2\mu'\Sigma^{-1}\mathbf{y}-2\mathbf{t}'\mathbf{y}+\mu'\Sigma^{-1}\mu=\mathbf{y}'\Sigma^{-1}\mathbf{y}-2(\mathbf{t}'\Sigma+\mu')\Sigma^{-1}\mathbf{y}+\mu'\Sigma^{-1}\mu$

$=[\mathbf{y}-(\Sigma\mathbf{t}+\mu)]'\Sigma^{-1}[\mathbf{y}-(\Sigma\mathbf{t}+\mu)]-\mu^{*\prime}\Sigma^{-1}\mu^*+\mu'\Sigma^{-1}\mu$

$M_\mathbf{Y}(\mathbf{t})=|\Sigma|^{-1/2}(2\pi)^{-p/2}\int\exp((\mathbf{y}-\mu^*)'\Sigma^{-1}(\mathbf{y}-\mu^*)/2)d\mathbf{y}\; e^{\mathbf{t}'\mu+\mathbf{t}'\Sigma\mathbf{t}/2}$

$=e^{\mathbf{t}'\mu+\mathbf{t}'\Sigma\mathbf{t}/2}$.

例 5.1

令 $\mathbf{y}=[Y_1,\ldots,Y_n]'\sim N(\alpha\mathbf{1},\sigma^2 I)$，$U=\sum(Y_i-\bar{y})^2/\sigma^2$, $V=n(\bar{y}-\alpha)^2/\sigma^2$,求 U 和 V 的分布，并证明 U 和 V 独立.

证明: $\bar{y}=\mathbf{1}'\mathbf{y}/n=A\mathbf{y}$, $(\mathbf{y}-\bar{y}\mathbf{1})=(I-\mathbf{1}\mathbf{1}'/n)\mathbf{y}=(I-J/n)\mathbf{y}=B\mathbf{y}$

B 是幂等的,因为$(I-J/n)^2=I-2J/n+J^2/n^2=I-J/n$, $rank(B)=tr(B)=n-1$.

$U=(\mathbf{y}-\bar{y}\mathbf{1})'(\mathbf{y}-\bar{y}\mathbf{1})/\sigma^2=\mathbf{y}'B'B\mathbf{y}/\sigma^2=\mathbf{y}'(B/\sigma^2)\mathbf{y}\sim\chi^2(n-1, \lambda)=\chi^2(n-1)$.

因为$(\bar{y}-\alpha)/(\sigma/\sqrt{n})\sim N(0,1)$，所以 $V\sim\chi^2(1)$.

因为 $A\Sigma(B/\sigma^2)=\mathbf{1}'/n\,\sigma^2 I\,(I-\mathbf{1}\mathbf{1}'/n)/\sigma^2=\mathbf{0}$，所以 U 和 $\bar{y}$ 独立，进而 U 和 V 独立，因为 V 是 $\bar{y}$ 的函数.

例 5.2 $\mathbf{Y}=[Y_1, \ldots, Y_n]'\sim N(\mu\mathbf{1}, \sigma^2 I)$, $Q_1=n(\bar{Y})^2$, $Q_2=\sum(Y_i-\bar{Y})^2$.

1.证明 $\bar{Y}$ 与 Q_2 独立；2.证明 Q_1 与 Q_2 独立；3.求 Q_1，Q_2 的分布.

解: $\bar{Y}=\mathbf{1}'\mathbf{y}/n=A\mathbf{y}$, $Q_2=(\mathbf{y}-\bar{y}\mathbf{1})'(\mathbf{y}-\bar{y}\mathbf{1})=\mathbf{y}'B'B\mathbf{y}=\mathbf{y}'B\mathbf{y}\sim\sigma^2\chi^2(n-1)$.

$A\Sigma B'=\mathbf{0}$，所以 $\bar{Y}$ 和 Q_2 独立，进而 Q_1 与 Q_2 也独立.

$Q_1=\mathbf{Y}'\mathbf{1}\mathbf{1}'\mathbf{Y}/n\sim\sigma^2\chi^2(1, n\mu^2/2\sigma^2)$.

例 5.3 $\mathbf{y}\sim N_3(\mu, \sigma^2 I)$, $\mu'=[3\ -2\ 1]'$, $A=[[2\ -1\ -1][-1\ 2\ -1][-1\ -1\ 2]]/3$

$B=[[1\ 1\ 1][1\ 0\ -1]]/3$.

求 $\mathbf{y}'A\mathbf{y}/\sigma^2$ 的分布；探究 $\mathbf{y}'A\mathbf{y}$ 与 $B\mathbf{y}$、$\mathbf{y}'A\mathbf{y}$ 与 $y_1+y_2+y_3$ 是否独立.

解: 易得$\sigma^2 AI$ 幂等，所以 $\mathbf{y}'A\mathbf{y}/\sigma^2\sim\chi^2(n-1, 38/3)$.

$\because B\sigma^2 IA\neq O$，所以 $\mathbf{y}'A\mathbf{y}$ 与 $B\mathbf{y}$ 不独立.

$y_1+y_2+y_3=[1\ 1\ 1]\mathbf{y}=\mathbf{C}'\mathbf{y}$, $\mathbf{C}'\sigma^2 IA=O$，所以 $\mathbf{y}'A\mathbf{y}$ 与 $y_1+y_2+y_3$ 独立.

例 5.4 $\mathbf{y}\sim N_n(\mu\mathbf{1}, \Sigma)$, $\Sigma=\sigma^2[[1\ \rho\ \ldots][\rho\ 1\ \ldots]\ \ldots\ [\ldots\ \rho\ 1]]$, 求

$U=\sum_{i=1}^n(y_i-\bar{y})^2/(\sigma^2(1-\rho))$的分布.

解: $\Sigma=\sigma^2((1-\rho)I+\rho J)$，记 $U=\mathbf{y}'A\mathbf{y}=\mathbf{y}'B\mathbf{y}/[\sigma^2(1-\rho)]$，则

$B=(I-J/n)$, $A=B/[\sigma^2(1-\rho)]$，而 $A\Sigma=B$，B 幂等，所以 $A\Sigma$幂等.

$rank(A)=rank(B)=tr(B)=n-1$, 所以 $U\sim\chi^2(n-1, \lambda)=\chi^2(n-1)$

Remark 3.3 证明 $X(X'X)^-X'$是不变的. X 是 m*n 的秩 k 矩阵.

证明: 由定义，$X'X(X'X)^-X'X=X'X$. 考虑满秩分解 $X=JL$，则有 $L'J'X(X'X)^-X'JL=L'J'JL$. 考虑 L 存在右逆 R，$LR=I_k$，易得 $J'X(X'X)^-X'J=J'J$，即 $J'JL(L'J'JL)^-L'J'J=J'J$，$L(L'J'JL)^-L'=(J'J)^{-1}$. 所以 $X(X'X)^-X'=JL(L'J'JL)^-L'J'=J(J'J)^{-1}J'$，这意味着其是不变的. 因为对于任意其他的满秩分解 $X=J^*L^*$，有 $J=J^*S_{k*k}$，其中 S 可逆. 所以 $J(J'J)^{-1}J'=J^*S(S'J^{*\prime}J^*S)^{-1}S'J^{*\prime}=J^*(J^{*\prime}J^*)^{-1}J^{*\prime}$.

Remark 4.5 $p(\mathbf{Y}_1|\mathbf{Y}_2=\mathbf{y}_2)=p(\mathbf{y}_1, \mathbf{y}_2)/p(\mathbf{y}_2)\propto p(\mathbf{y}_1, \mathbf{y}_2)$

$p(\mathbf{y}_1, \mathbf{y}_2)\propto\exp(-[[\mathbf{y}_1-\mu_1]\ [\mathbf{y}_2-\mu_2]]'\Sigma^{-1}[[\mathbf{y}_1-\mu_1]\ [\mathbf{y}_2-\mu_2]]/2)\propto\exp(((\mathbf{y}_1-\mu_1)'B^{-1}(\mathbf{y}_1-\mu_1)-2(\mathbf{y}_1-\mu_1)'B\Sigma_{12}\Sigma_{22}^{-1}(\mathbf{y}_2-\mu_2)+(\mathbf{y}_2-\mu_2)'A(\mathbf{y}_2-\mu_2))/2)$

$\propto\exp(-(\mathbf{y}_1-\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(\mathbf{y}_2-\mu_2))'B^{-1}(\mathbf{y}_1-\mu_1-\Sigma_{12}\Sigma_{22}^{-1}(\mathbf{y}_2-\mu_2))/2)$

$\mathbf{Y}_1|\mathbf{Y}_2=\mathbf{y}_2\sim N(\mu_1+\Sigma_{12}\Sigma_{22}^{-1}(\mathbf{y}_2-\mu_2), \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$

其中 $B=Cov(\mathbf{Y}_1|\mathbf{Y}_2)=\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

Remark 4.6 Correlation and Partial Correlation

y-血压; x_1-摄入热量; x_2-体重. $\rho_1=corr(y, x_1)$, $\rho_2=corr(y, x_2)$.

$\rho_3=corr(x_1, x_2)\neq0$. 如何将 x_2 加入线性模型 $y=a+bx_1+\varepsilon_1$?

假设 x_1 与 x_2 之间也有线性关系 $x_2=c+dx_1+\varepsilon_2$,

记残差 $r_y=y-(a^{hat}+b^{hat}x_1)$，$r_2=x_2-(c^{hat}+d^{hat}x_1)$，则

$\rho_(y, x_2).x_1=corr(r_y, r_2)$为给定 x_1 时，x_2 和 y 的偏相关系数.

偏相关为 0 时，意味着 x_2 中能解释 y 的信息均包含在 x_1 中，因此无需考虑将 x_2 加入模型；而偏相关不为 0 时，意味着 x_2 包含 x_1 以外的能够解释 y 的额外信息，此时应考虑将 x_2 加入模型.

Remark 5.1 二次型矩生成函数的表达式推导

$M_{\mathbf{x}'A\mathbf{x}}(t)=E(e^{t\mathbf{x}'A\mathbf{x}})=c_1\int\ldots\int\exp(t\mathbf{x}'A\mathbf{x}-(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)/2)d\mathbf{x}$

$=c_1\int\ldots\int\exp(*)d\mathbf{x}$. 其中，$c_1=(2\pi)^{-p/2}|\Sigma|^{-1/2}$.

记 $V^{-1}=(I-2tA\Sigma)\Sigma^{-1}$，$\theta'=\mu'(I-2tA\Sigma)^{-1}$，则

$(*)=-[\mathbf{x}(I-2tA\Sigma)^{-1}\Sigma^{-1}\mathbf{x}-2\mu'\Sigma^{-1}\mathbf{x}+\mu'\Sigma^{-1}\mu]/2$

$=-(\mathbf{x}-\theta)'V^{-1}(\mathbf{x}-\theta)/2-(\mu'\Sigma^{-1}\mu-\theta'V^{-1}\theta)/2$，于是

$M_{\mathbf{x}'A\mathbf{x}}(t)=(|V|/|\Sigma|)^{-1/2}\exp(-(\mu'\Sigma^{-1}\mu-\theta'V^{-1}\theta)/2)$，其中

$\theta'V^{-1}\theta=\mu'\Sigma^{-1}(I-2t\Sigma A)^{-1}\mu=\mu'\Sigma^{-1}(\Sigma(I-2tA\Sigma)^{-1}\Sigma^{-1})^{-1}\mu$

$=\mu'(I-2tA\Sigma)^{-1}\Sigma^{-1}\mu$. 所以

$M_{\mathbf{x}'A\mathbf{x}}(t)=|I-2tA\Sigma|^{-1/2}\exp(-\mu'(I-(I-2tA\Sigma)^{-1})\Sigma^{-1}\mu/2)$.

Remark 5.2 用矩生成函数推导二次型的期望和方差

$E(y^n)=d^nM_y(t)/dt^n|_{t=0}$，$Var(y)=\partial\ln M_y(t)/\partial t^2|_{t=0}$.

期望的证明：见 Remark 4.2

方差的证明：

令 $c(t)=I-2tA\Sigma$，则 $c(0)=I$. 令 $k(t)=\ln M_{\mathbf{x}'A\mathbf{x}}(t)$，则

$k(t)=-(\ln|c(t)|)/2-\mu'(I-c^{-1}(t))\Sigma^{-1}\mu$.

$\partial k(t)/\partial t=-tr(c^{-1}(t)\,\partial c(t)/\partial t)/2 + \mu'\,\partial c^{-1}(t)/\partial t\,\Sigma^{-1}\mu/2$

$=tr(c^{-1}(t)A\Sigma)+\mu'c^{-1}(t)A\Sigma c^{-1}(t)\Sigma^{-1}\mu$

$\partial k^2(t)/\partial t^2=tr(\partial c^{-1}(t)/\partial t\, A\Sigma)+\mu'[\partial c^{-1}(t)/\partial t]A\Sigma[c^{-1}(t)]\Sigma^{-1}\mu$

$+\mu'[c^{-1}(t)]A\Sigma[\partial c^{-1}(t)/\partial t]\Sigma^{-1}\mu$

$=2tr(c^{-1}(t)A\Sigma c^{-1}(t)A\Sigma)+4\mu'c^{-1}(t)A\Sigma c^{-1}(t)A\Sigma c^{-1}(t)\Sigma^{-1}\mu$

$Var(\mathbf{x}'A\mathbf{x})=\partial^2k(t)/\partial t^2|_{t=0}=2tr((A\Sigma)^2)+4\mu'A\Sigma A\mu$

If A, B and A+PBQ are nonsingular, then
$(A+PBQ)^{-1}=A^{-1}-A^{-1}PB(B+BQA^{-1}PB)^{-1}BQA^{-1}$

例 证明$(X'X)^{-}X'$对于任意的$(X'X)^{-}$都是 X 的广义逆.
证明：由定义，$X'X(X'X)^{-}X'X=X'X$
即 $X'X((X'X)^{-}X'X-I)=0$，所以
$((X'X)^{-}X'X-I)'X'X((X'X)^{-}X'X-I)=0$，
$(X((X'X)^{-}X'X-I))'X((X'X)^{-}X'X-I)=0$.
$X((X'X)^{-}X'X-I)=X((X'X)^{-}X')X=X$，这正是 X 的广义逆的定义.

例 $\mathbf{X}\sim N_3(\boldsymbol{\mu}, \Sigma)$，$\boldsymbol{\mu}'=[3\ {-2}\ 0]$，$\mathbf{X}'=[X_1\ X_2\ X_3]$，
$\Sigma=[[5\ 0\ {-3}]$
$[0\ 9\ 0]$
$[-3\ 0\ 2]]$
a. X_2 和 $2X_1-X_3$ 是否独立?
b. 找到$(2X_1-5X_3, X_1+X_2)$的分布.
c. 给定 $X_1=1$，$X_2=-2$，求 X_3 的条件分布.
解：a.令 $B=[[0\ 1\ 0][2\ 0\ {-1}]]$，则 $B\mathbf{X}=[[X_2][2X_1-X_3]]$.
$B\mathbf{X}\sim N(B\boldsymbol{\mu}, B\Sigma B')=N([-2\ 6]', [[9\ 0][0\ 34]])$.
注意到 $Cov(X_2, 2X_1-X_3=0$，所以它们独立.
b. 令 $C=[[2\ 0\ {-5}][1\ 1\ 0]]$，则 $B\mathbf{X}=[[2X_1-5X_3][X_1+X_2]]$.
$B\mathbf{X}\sim N([6\ 1]', [[130\ 25][25\ 14]])$.
c. $X_3|X_1=1, X_2=2 \sim N(0+[-3\ {-0}][[5\ 0][0\ 9]]^{-1}([1\ 2]-[3\ {-2}])'$,
$[[2]]-[-3\ {-0}][[5\ 0][0\ 9]]^{-1}[-3\ {-0}]')$
$=N((6/5, 1/5)$.

例 有下列数据

x:	4.4	3.9	4.0	4.0	3.5	4.1
y:	78	74	68	76	73	84

I. 求最小二乘回归直线.
II. 你的回归模型是什么？陈述必要的假设.
III. 检验当使用线性模型时，y 与 x 无影响这一假设. 陈述零假设和备择假设，并导出测试结论. ($\alpha=0.05$)
IV. 求β_1 的 95%置信区间. 解释你的结果.
V. 求模型的 Coefficient of Determination. 解释你的结果.
VI. 对于一个新的值 x=4，求 y 的预测值的 95%置信区间.
解：[菜单] [6] [2] [4.4 = 3.9 = 4.0 = 4.0 = 3.5 = 4.1][→][↑]
[78 = 74 = 68 = 76 = 73 = 84]
-[OPTN][3] ->
$x^{-}=3.983$; $\sum x=23.9$; $\sum x^2=95.63$; $\sigma_x^2=0.0714$; $S_x^2=0.0856$
$y^{-}=75.5$; $\sum y=453$; $\sum y^2=34345$; $\sigma_y^2=23.91$; $S_y^2=28.7$
$\sum xy=1807.7$;   [OPTN]
-[OPTN][4] ->
y=ax+b: a=7.586; b=45.27; r=0.4145.

(!):
$S_{xx} = \sum x^2-(\sum x)^2/n= n\sigma_x^2 =95.63-(23.9)^2/6=0.4283$;
$S_{xy} = \sum xy-\sum x\sum y/n=1807.7-23.9*453/6=3.25$
$SSE = n\sigma_y^2(1-r^2)=118.81$;
$SST=n\sigma_y^2=143.5$; $SSR=n\sigma_y^2r^2=24.66$
$S^2=SSE/(n-2)=29.703$

I. $y^{hat}=45.27+7.586x$
II. $y_i=\beta_0+\beta_1x_i+\varepsilon_i$，并假定 y 与 x 之间的关系是线性的，
$\varepsilon_i\sim N(0, \sigma^2)$，且$\varepsilon_i$ 相互独立.
III & IV. 零假设：$\beta_1=0$ 备择假设：$\beta_1\neq0$.
检验统计量 $t=(\beta_1^{hat}-\beta_1)(\sqrt{S_{xx}})/\sqrt{S^2}\sim t(n-2)$
95%置信区间$[\beta_1^{hat}\pm S\ t(0.025, n-2)/\sqrt{S_{xx}}]=[-15.5, 30.7]$
因为 $0\in[-15.5, 30.7]$故不能拒绝零假设. 这意味着我们没有95%的把握认为 x 和 y 具有线性相关关系.
V. $R^2=SSR/SST=1-SSE/SST=r^2=0.1718$. 这说明 x 和 y 的关系中仅有 0.1718 的部分可以被模型解释.
VI. $y_h=45.27+7.586*4.0=75.6265$.
95%置信区间$[y_h\pm t_{\alpha/2, n-2}S\sqrt{(1+1/n+(x_h-x^{-})^2/\sum_i(x_i-x^{-})^2)}]$
$=[59.2758, 91.9771]$.