Biostat 250B • HW5

Elvis Cui
Han Cui
Dept. of Biostat
UCLA

1 Let X be a chi-squared distribution with n degrees of freedom. Find $EX^{1/2}$ and hence find an unbiased estimate of σ in the standard linear model EY=Xβ where X is n x p with full column rank and error terms are independent, each with mean 0 and variance σ^2.

2 Let μ belong to V, a p-dimensional vector space of R^n, let y belong to R^n and let P_V be the orthogonal projection matrix onto V.

If $\hat{\mu} = P_V y$ and $\hat{\sigma}^2 = ||(I-P_V)y||^2/(n-p)$ show that

i) $E||\hat{\mu}||^2$ always overestimates $||\mu||^2$,

ii) $E||\hat{\mu} - \mu||^2/\sigma^2 = p$ and identify the distribution of $||\hat{\mu} - \mu||^2$,

iii) $E\hat{\sigma}^4 = (n-p+2)\sigma^4/(n-p)$.

3 Recall that if X and Y are random variables with finite means and variances EEX|Y = EX. Use this result to show that if the conditional distribution of V given K is a central chi-squared with p+2K degrees of freedom, then the conditional expectation of 1/V given K is 1/(p-2+2K).

4 Suppose the conditional density of y given μ is univariate normal with mean μ and variance σ^2. If μ is univariate normal with mean μ_0 and variance σ^{*2}, find the conditional density of μ given y. Can you generalize the above setting when you now have a random sample size $y_1, y_2, .., y_n$? Express the conditional density of μ given $y_1, y_2, .., y_n$ in terms of n, μ_0, σ^2 and σ^{*2}. [These resulting distributions are called posterior distributions].

5 To obtain Bayesian estimates for the vector of parameters β in the standard linear model, answers to this and the next question are helpful. First recall the gamma function defined by $\Gamma(x) = \int_0^\infty t^{x-1}\exp(-t)dt$.

Show that $\int_0^\infty \exp(-k/x)x^{-v-1}dx = k^{-v}\Gamma(v)$ and $\int_0^\infty \exp(-a/x^2)x^{-b-1}dx = \frac{1}{2}a^{-b/2}\Gamma(b/2)$.

6 Recall the result from Biostat 250A that if we have conformable matrices A, B, U and V, and the indicated inverses all exist, we have

$$(A+UBV)^{-1} = A^{-1} - A^{-1}UB(B + BVA^{-1}UB)^{-1}BVA^{-1}.$$

Use the above result to show that

$$(y-X\beta)'(y-X\beta) + (\beta-m)'V^{-1}(\beta-m) = (\beta-m^*)'V^{*-1}(\beta-m^*) + (y - Xm)'(I + XVX')^{-1}(y - Xm),$$

where $V^* = (X'X+V^{-1})^{-1}$ and $m^* = V^*(X'y + V^{-1}m)$.

[This is the essence of Theorem 3.7 in Biostat 250A textbook. I urge you to try hard to do the algebra on your own before referring to the book on pages 74-76 that uses much of what we learned from last quarter.]

Q1: $X \sim \chi^2_n(0)$

(a) The pdf of X is

$$f_X(x) = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$$

$$\mathbb{E} X^{\frac{1}{2}} = \int_{\mathbb{R}^+} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n+1}{2}-1} e^{-\frac{x}{2}} dx$$

$$= \sqrt{2}\, \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}.$$

(b): We know that

$$\hat{\sigma}^2 = \frac{Y^T Q_X Y}{n-p} \text{ is unbiased for } \sigma^2$$

where $Q_X = I - X(X^TX)^{-1}X^T$.

& $(n-p)\hat{\sigma}^2/\sigma^2 \sim \chi^2_{n-p}$ by Cochran,

thus,

$$\mathbb{E}\hat{\sigma} = \sqrt{\frac{\sigma^2}{n-p}}\, \mathbb{E} Z^{\frac{1}{2}} \text{ where } Z \sim \chi^2_{n-p}$$

$$\Rightarrow \mathbb{E}\hat{\sigma} = \frac{\sigma}{\sqrt{n-p}} \sqrt{2}\, \frac{\Gamma(\frac{n-p+1}{2})}{\Gamma(\frac{n-p}{2})}$$

Thus, an unbiased estimator for σ:

$$\tilde{\sigma} = \sqrt{\frac{n-p}{2}}\, \frac{\Gamma(\frac{n-p}{2})}{\Gamma(\frac{n-p+1}{2})}\, \hat{\sigma} \quad \text{where}$$

$$\hat{\sigma} = \left(\frac{Y^T Q_X Y}{n-p}\right)^{\frac{1}{2}}$$

Q2:

2 Let μ belong to V, a p-dimensional vector space of R^n, let y belong to R^n and let P_V be the orthogonal projection matrix onto V.

If $\hat{\mu} = P_V y$ and $\hat{\sigma}^2 = ||(I-P_V)y||^2/(n-p)$ show that

i) $E||\hat{\mu}||^2$ always overestimates $||\mu||^2$,

ii) $E||\hat{\mu} - \mu||^2/\sigma^2 = p$ and identify the distribution of $||\hat{\mu} - \mu||^2$,

iii) $E\hat{\sigma}^4 = (n-p+2)\sigma^4/(n-p)$.

Sol. (i): Let $f(x) = x^Tx$. then

$$\frac{\partial^2 f}{\partial x \partial x^T} = 2I \succ 0 \Rightarrow f(x) \text{ is a cvx function.}$$

$$\Rightarrow \mathbb{E}\|\hat{\mu}\|_2^2 = \mathbb{E} f(\hat{\mu}) \overset{(\Delta)}{\geq} f(\mathbb{E}\hat{\mu})$$

$$= f(\mathbb{E}(P_V Y)) \overset{(*)}{=} f(\mu) = \mu^T\mu.$$

(Δ): Jensen's ineq.

$(*)$: $\mu \in V \Rightarrow \mathbb{E} P_V Y = P_V \mathbb{E} Y = P_V \mu = \mu$.

(ii): $\mathbb{E}\|\hat{\mu} - \mu\|_2^2 = \mathbb{E}\|P_V Y - P_V \mu\|_2^2$

$$= \mathbb{E}(Y-\mu)^T P_V (Y-\mu)$$

$$= \mathbb{E}\, \mathrm{Tr}(P_V (Y-\mu)(Y-\mu)^T)$$

$$= \mathrm{Tr}\, P_V \underbrace{(\mathbb{E}(Y-\mu)(Y-\mu)^T)}_{Cov(Y)}$$

$$= \mathrm{Tr}\, P_V (\sigma^2 I)$$

$$= \sigma^2\, \mathrm{Tr}(P_V) \overset{(\square)}{=} \sigma^2 p$$

$(\square)$: $\mathrm{Tr}(P_V) = \mathrm{rank}(P_V) = \dim(V) = p$.

$$\Rightarrow \mathbb{E}\|\hat{\mu} - \mu\|_2^2/\sigma^2 = p.$$

Moreover, we have (Since $\mu \in V, \mu \perp V^\perp$)

$$(Y-\mu)^T(Y-\mu) = (Y-\mu)^T P_V (Y-\mu) + (Y-\mu)^T(I-P_V)(Y-\mu)$$

& $(Y-\mu)^T(Y-\mu) \sim \sigma^2\chi^2_n$

By Cochran, $\mathrm{rank}(P_V) + \mathrm{rank}(I-P_V) = \mathrm{Tr}(P_V) + \mathrm{Tr}(I-P_V) = n$ implies

$$\frac{(Y-\mu)^T P_V (Y-\mu)}{\sigma^2} \sim \chi^2_m,$$

$m = \mathrm{rank}(P_V) = \mathrm{Tr}(P_V) = p$.

But $P_V(Y-\mu) = \hat{\mu} - \mu \Rightarrow$

$$\|\hat{\mu} - \mu\|_2^2 \sim \chi^2_p \sigma^2$$

Q2 Cont'd:

(iii) $\hat{\sigma}^2 = \|(I-P_V)Y\|_2^2/(n-p)$

By (ii), we know that

$\frac{1}{\sigma^2}(Y-\mu)^T(I-P_V)(Y-\mu) \sim \chi^2_{n-p}$

by Fisher-Cochran's theorem.

But $\|(I-P_V)Y\|_2^2 = (Y-\mu)^T(I-P_V)(Y-\mu)$ $\Rightarrow$

Since $(I-P_V)\mu = \mu-\mu = 0$.

$\|(I-P_V)Y\|_2^2 \sim \sigma^2\chi^2_{n-p}$ $\Rightarrow$

$$\mathbb{E}\|(I-P_V)Y\|_2^4 = \sigma^4\big(\underbrace{2(n-p)}_{Var(\cdot)} + \underbrace{(n-p)^2}_{(\mathbb{E}(\cdot))^2}\big)$$

$$= \sigma^4(n-p)(n+2-p) \Rightarrow$$

$$\mathbb{E}\,\hat{\sigma}^4 = \frac{n-p+2}{n-p}\sigma^4 \qquad \square.$$

Q3:

3 Recall that if X and Y are random variables with finite means and variances EEX|Y = EX. Use this result to show that if the conditional distribution of V given K is a central chi-squared with p+2K degrees of freedom, then the conditional expectation of 1/V given K is 1/(p-2+2K).

Sol. $V|K \sim \chi^2_{p+2K}(0)$, then

Conditional density of $V|K$ is

$$f_{V|K}(v) = \frac{1}{2^{\frac{p+2k}{2}}\Gamma(\frac{p+2k}{2})} v^{\frac{p+2k}{2}-1} e^{-\frac{v}{2}}$$

& we have

$$\mathbb{E}\left(\tfrac{1}{V}\middle|K\right) = \int_{\mathbb{R}^+} \frac{1}{v} \mathbb{P}_V(dv)$$

$$= \int_{\mathbb{R}^+} \frac{1}{2^{\frac{p+2k}{2}}\Gamma(\frac{p+2k}{2})} v^{\frac{p+2k-2}{2}-1} e^{-\frac{v}{2}} dv$$

$$= \frac{1}{2}\Gamma\left(\frac{p+2k}{2}-1\right)\Big/\Gamma\left(\frac{p+2k}{2}\right)$$

$$\overset{(*)}{=} \frac{1}{p+2k-2}$$

(*): $\Gamma(n) = n\Gamma(n-1)$

Q4:

4 Suppose the conditional density of y given μ is univariate normal with mean μ and variance σ^2. If μ is univariate normal with mean μ_0 and variance σ^{*2}, find the conditional density of μ given y. Can you generalize the above setting when you now have a random sample size $y_1, y_2, \ldots, y_n$? Express the conditional density of μ given $y_1, y_2, \ldots, y_n$ in terms of n, μ_0, σ^2 and σ^{*2}. [These resulting distributions are called posterior distributions].

Sol. (a): $f(\mu|y) \propto f(y|\mu)f(\mu)$

& $f(y|\mu)f(\mu)$

$= C e^{-\frac{(y-\mu)^2}{2\sigma^2}} e^{-\frac{(\mu-\mu_0)^2}{2\sigma^{*2}}}$, C is a constant

$= g(y)\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2}+\frac{1}{\sigma^{*2}}\right)\mu^2 + \left(\frac{y}{\sigma^2}+\frac{\mu_0}{\sigma^{*2}}\right)\mu\right)$

where $g(y)$ is a function of y

Thus, we know that $\mu|y$ is also normally distributed (normal kernel) & the varince is $\left(\frac{1}{\sigma^2}+\frac{1}{\sigma^{*2}}\right)^{-1}$, the mean is $\left(\frac{y}{\sigma^2}+\frac{\mu_0}{\sigma^{*2}}\right)/\left(\frac{1}{\sigma^2}+\frac{1}{\sigma^{*2}}\right)$ or

$$\frac{\sigma^{*2}}{\sigma^2+\sigma^{*2}}y + \frac{\sigma^2}{\sigma^2+\sigma^{*2}}\mu_0$$

(b): $f(\mu|y_1,\ldots,y_n) \propto \prod_{i=1}^{n} f(y_i|\mu)f(\mu)$.

Note $\prod_{i=1}^{n} f(y_i|\mu)$

$= C\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right)$, C is a constant

$= C\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\bar{y}+\bar{y}-\mu)^2\right)$, $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$

$= g(y)\exp\left(-\frac{n}{2\sigma^2}(\bar{y}-\mu)^2\right)$,

$g(y)$ is a fct of $(y_1,\ldots,y_n)$.

Thus, by result from (a), we have

$\mu|y_1,\ldots,y_n \sim \mathcal{N}(a, b)$,

$b = \left(\frac{n}{\sigma^2}+\frac{1}{\sigma^{*2}}\right)^{-1}$ and

$a = \frac{n\sigma^{*2}}{\sigma^2+n\sigma^{*2}}\bar{y} + \frac{\sigma^2}{\sigma^2+n\sigma^{*2}}\mu_0$.

Q5:

5 To obtain Bayesian estimates for the vector of parameters β in the standard linear model, answers to this and the next question are helpful. First recall the gamma function defined by $\Gamma(x) = \int_0^\infty t^{x-1}\exp(-t)dt$. Show that $\int_0^\infty \exp(-k/x)x^{-v-1}dx = k^{-v}\Gamma(v)$ and $\int_0^\infty \exp(-a/x^2)x^{-b-1}dx = \frac{1}{2}a^{-b/2}\Gamma(b/2)$.

Sol. $\int_0^\infty e^{(-\frac{k}{x})} x^{-v-1}dx$

$= \left[\int_0^\infty e^{(-\frac{k}{x})}\left(\frac{k}{x}\right)^{v-1} d\left(\frac{k}{x}\right)\right]k^{-v}$ Change of variable formula

$= k^{-v}\Gamma(v)$

- $-d\frac{a}{x^2} = \frac{2a}{x^3}dx$, thus,

$\int_0^\infty e^{-\frac{a}{x^2}} x^{-b-1}dx$

$= \int_0^\infty e^{-\frac{a}{x^2}} \cdot x^{-b-1} \cdot \frac{x^3}{2a} d\frac{a}{x^2}$ change of variable

$= \left[\int_0^\infty e^{-\frac{a}{x^2}}\left(\frac{a}{x^2}\right)^{\frac{b}{2}-1} d\frac{a}{x^2}\right]\left(\frac{1}{2}a^{-\frac{b}{2}}\right)$

$= \frac{1}{2}a^{-\frac{b}{2}}\Gamma\left(\frac{b}{2}\right)$

Q6:

6 Recall the result from Biostat 250A that if we have conformable matrices A, B, U and V, and the indicated inverses all exist, we have

$$(A+UBV)^{-1} = A^{-1} - A^{-1}UB(B + BVA^{-1}UB)^{-1}BVA^{-1}.$$

Use the above result to show that

$$(y-X\beta)'(y-X\beta) + (\beta-m)'V^{-1}(\beta-m) = (\beta-m^*)'V^{*-1}(\beta-m^*) + (y - Xm)'(I + XVX')^{-1}(y - Xm),$$

where $V^* = (X'X+V^{-1})^{-1}$ and $m^* = V^*(X'y + V^{-1}m)$.

[This is the essence of Theorem 3.7 in Biostat 250A textbook. I urge you to try hard to do the algebra on your own before referring to the book on pages 74-76 that uses much of what we learned from last quarter.]

Sol.

$$\begin{aligned} LHS &= y^Ty - 2y^TX\beta + \beta^TX^TX\beta + \beta^TV^{-1}\beta - 2m^TV^{-1}\beta + m^TV^{-1}m \\ &= \beta^T(X^TX+V^{-1})\beta - 2(X^Ty+V^{-1}m)^T\beta + y^Ty + m^TV^{-1}m \end{aligned}$$

By completing the square,

$$\beta^T(X^TX+V^{-1})\beta - 2(X^Ty+V^{-1}m)^T\beta = (\beta-m^*)^TV^{*-1}(\beta-m^*) - m^{*T}V^{*-1}m^*$$

(The above is part I)

And

$$(m^*)^T(V^*)^{-1}(m^*) = (X^Ty+V^{-1}m)^TV^*(X^Ty+V^{-1}m) \quad (\Delta)$$

Also by 250A,

$$\begin{aligned}(I+XVX^T)^{-1} &= I - X(V^{-1}+X^TX)^{-1}X^T \\ &= I - XV^*X^T\end{aligned}$$

Moreover,

$$XV^*V^{-1} = (I - XV^*X^T)X.$$

$$V^*V^{-1} = I - V^*X^TX$$

Thus, we have

$$-m^{*T}V^{*-1}m^* + y^Ty + m^TV^{-1}m = (y-Xm)^T(I-XV^*X^T)(y-Xm).$$

To sum up,

$$(y-X\beta)^T(y-X\beta) + (\beta-m)^TV^{-1}(\beta-m) = (\beta-m^*)^TV^{*-1}(\beta-m^*) + (y-Xm)^T(I+XVX^T)^{-1}(y-Xm).$$