24500 HW3

Bin Yu

January 28, 2025

Question 1

(a)

The joint PDF of $X_1, \ldots, X_n$ for a Normal distribution $N(\mu, \sigma^2)$ is

$$f\left(x_1, \ldots, x_n \mid \mu, \sigma^2\right) = \frac{1}{(2\pi\,\sigma^2)^{\frac{n}{2}}} \exp\Bigl[-\tfrac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\Bigr].$$

When μ is unknown, we can use the MLE for μ for substitution.

$$\ln L(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi\,\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2.$$

take the derivative of $\ln L(\mu, \sigma^2)$ with respect to μ and set it to zero:

$$\frac{\partial}{\partial \mu} \ln L(\mu, \sigma^2) = \frac{\partial}{\partial \mu} \left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right).$$

$$\frac{\partial}{\partial \mu} \ln L(\mu, \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu).$$

$$\frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0.$$

$$\sum_{i=1}^{n} (x_i - \mu) = 0.$$

$$\sum_{i=1}^{n} x_i - n\mu = 0.$$

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

Plug $\mu = \overline{x}$ back into the likelihood, get

$$L(\sigma^2) = \frac{1}{(2\pi\,\sigma^2)^{\frac{n}{2}}} \exp\Bigl[-\tfrac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \overline{x})^2\Bigr].$$

Taking log,

$$\ell(\sigma^2) = -\frac{n}{2} \ln(2\pi\,\sigma^2) \;-\; \frac{1}{2\,\sigma^2} \sum_{i=1}^{n} (x_i - \overline{x})^2.$$

$$\ell(\sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\sigma^2 \; - \; \frac{1}{2\,\sigma^2}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Differentiate on σ^2, set equal to 0:

$$\frac{\partial \ell}{\partial(\sigma^2)} = -\frac{n}{2}\,\frac{1}{\sigma^2} \; + \; \frac{1}{2\,(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \overline{x})^2 = 0,$$

$$\frac{1}{2\,(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{n}{2\,\sigma^2}$$

$$\hat{\sigma}^2_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Hence the MLE corresponds to $c = \frac{1}{n}$.

$$\hat{\sigma}^2_{\text{MLE}} = \frac{1}{n}\,\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

$$c = \frac{1}{n}$$

(b)

By definition,

$$\hat{\sigma}^2_c = c\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

Since: $\sum_{i=1}^{n}(X_i - \overline{X})^2/\sigma^2 \sim \chi^2_{n-1}$, and

$$E[\chi^2_n] \;=\; n \quad \text{and} \quad \mathrm{Var}(\chi^2_n) \;=\; 2n.$$

$$E\Big[\sum_{i=1}^{n}(X_i - \overline{X})^2\Big] = E\Big[\chi^2_{n-1}\sigma^2\Big] = \sigma^2 E\Big[\chi^2_{n-1}\Big] = (n-1)\,\sigma^2$$

$$\mathrm{Var}\Big[\sum_{i=1}^{n}(X_i - \overline{X})^2\Big] = \mathrm{Var}[\chi^2_{n-1}\sigma^2] = \sigma^4\mathrm{Var}[\chi^2_{n-1}] = 2\,(n-1)\,\sigma^4.$$

Therefore,

$$E\big[\hat{\sigma}^2_c\big] = c\,(n-1)\,\sigma^2, \quad \mathrm{Var}\big[\hat{\sigma}^2_c\big] = c^2 \cdot 2\,(n-1)\,\sigma^4.$$

For $\hat{\sigma}^2_c$ to be unbiased, we need

$$E\big[\hat{\sigma}^2_c\big] = \sigma^2$$

$$c\,(n-1)\,\sigma^2 = \sigma^2$$

$$c = \frac{1}{n-1}.$$

(c)

$$\mathrm{MSE}\big[\hat{\sigma}^2_c\big] = E\big[(\hat{\sigma}^2_c - \sigma^2)^2\big].$$

can be decomposed as

$$\mathrm{MSE}\big[\hat{\sigma}^2_c\big] = \mathrm{Var}\big[\hat{\sigma}^2_c\big] \; + \; \big(E[\hat{\sigma}^2_c] - \sigma^2\big)^2.$$

From above, we have

$$\mathrm{Var}\big[\hat{\sigma}^2_c\big] = 2\,c^2\,(n-1)\,\sigma^4, \quad E[\hat{\sigma}^2_c] - \sigma^2 = c\,(n-1)\,\sigma^2 - \sigma^2 = \sigma^2\big[c(n-1) - 1\big].$$

Hence

$$\mathrm{MSE}\big[\hat{\sigma}_c^2\big] = 2\,c^2\,(n-1)\,\sigma^4 \;+\; \sigma^4\big[c(n-1)-1\big]^2.$$

$$\mathrm{MSE}\big[\hat{\sigma}_c^2\big] = \sigma^4\Big\{2\,c^2\,(n-1) + \big[c(n-1)-1\big]^2\Big\}.$$

Since σ^4 is always positive, to find the minimizing c, treat the bracketed expression as a function of c and set its derivative to zero. Let

$$g(c) = 2\,c^2\,(n-1) + \big[c(n-1)-1\big]^2.$$

$$g(c) = 2(n-1)\,c^2 + (n-1)^2\,c^2 - 2(n-1)\,c + 1.$$

$$g(c) = \big[2(n-1) + (n-1)^2\big]c^2 - 2(n-1)\,c + 1.$$

$$g(c) = (n-1)\big[2+(n-1)\big]c^2 - 2(n-1)\,c + 1 = (n-1)(n+1)\,c^2 - 2(n-1)\,c + 1.$$

Differentiate over c:

$$g'(c) = 2(n-1)(n+1)\,c - 2(n-1).$$

Set $g'(c) = 0$:

$$2(n-1)(n+1)\,c - 2(n-1) = 0$$

$$(n+1)\,c - 1 = 0.$$

Thus the optimal choice of c that minimizes MSE is

$$c = \frac{1}{n+1}.$$

Question 2

Set

$$X = Z_1, \quad Y = Z_2,$$

and consider the transformation:

$$U = \frac{X}{Y}, \quad V = Y.$$

Then we can solve for X and Y in terms of U and V:

$$X = U\,V, \quad Y = V.$$

the Jacobian determinant of $(u,v) \mapsto (x,y)$:

$$\frac{\partial(x,y)}{\partial(u,v)} \;=\; \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \;=\; \begin{vmatrix} v & u \\ 0 & 1 \end{vmatrix} \;=\; v.$$

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = |v|.$$

Because X and Y are independent $N(0,1)$ random variables, their joint density is

$$f_{X,Y}(x,y) \;=\; \frac{1}{2\pi}\,\exp\!\Big(-\tfrac{x^2+y^2}{2}\Big).$$

Using the transformation $x = uv$, $y = v$ and the Jacobian, we get

$$f_{U,V}(u,v) \;=\; f_{X,Y}(uv,\,v) \;\times\; \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \;=\; \frac{1}{2\pi}\,\exp\!\Big(-\tfrac{(uv)^2+v^2}{2}\Big)\,|v|.$$

$$f_{U,V}(u,v) \;=\; \frac{1}{2\pi}\,\exp\!\Big(-\tfrac{v^2\,(u^2+1)}{2}\Big)\,|v|.$$

To find the distribution of $U = X/Y$, integrate out v:

$$f_U(u) \;=\; \int_{-\infty}^{\infty} f_{U,V}(u,v)\,\mathrm{d}v \;=\; \int_{-\infty}^{\infty} \frac{1}{2\pi}\,\exp\!\Big(-\tfrac{v^2\,(u^2+1)}{2}\Big)\,|v|\,\mathrm{d}v.$$

Since the integrand is even in v (because $|v|$ is an even function and the exponential depends on v^2), we can convert the integral from $-\infty$ to ∞ into twice the integral from 0 to ∞. The factor of $\frac{1}{2\pi}$ then becomes $\frac{1}{\pi}$:

$$f_U(u) \;=\; \frac{1}{\pi}\int_0^{\infty} \exp\!\Big(-\tfrac{v^2\,(u^2+1)}{2}\Big)\,v\,\mathrm{d}v.$$

To simplify, make the substitution:

$$t \;=\; \frac{(u^2+1)\,v^2}{2} \quad\Longrightarrow\quad \mathrm{d}t \;=\; (u^2+1)\,v\,\mathrm{d}v, \quad v\,\mathrm{d}v \;=\; \frac{\mathrm{d}t}{u^2+1}.$$

When v goes from 0 to ∞, so does t:

$$f_U(u) \;=\; \frac{1}{\pi}\int_0^{\infty} e^{-t}\,\frac{\mathrm{d}t}{u^2+1} \;=\; \frac{1}{\pi\,(u^2+1)}\int_0^{\infty} e^{-t}\,\mathrm{d}t.$$

The integral $\int_0^\infty e^{-t}\,\mathrm{d}t$ equals 1, so:

$$f_U(u) \;=\; \frac{1}{\pi\,(u^2+1)}.$$

This is the probability density function of Cauchy$(0,1)$. Thus:

$$U = \frac{Z_1}{Z_2} \sim \text{Cauchy}(0,1)$$

Question 3

By definition,

$$\mathrm{Cov}(AX,\; BY) \;=\; E\Big[\big(AX - E[AX]\big)\,\big(BY - E[BY]\big)^{\top}\Big].$$

We have $E[AX] = A\,E[X]$ and $E[BY] = B\,E[Y]$ by linearity of expectation. Thus:

$$AX - E[AX] \;=\; A\,(X - E[X]), \quad BY - E[BY] \;=\; B\,(Y - E[Y]).$$

$$\mathrm{Cov}(AX, BY) = E\Big[A\,(X - E[X])\,\big(B\,(Y - E[Y])\big)^{\top}\Big].$$

By definition:

$$\big(B\,(Y - E[Y])\big)^{\top} \;=\; (Y - E[Y])^{\top}\,B^{\top}.$$

$$\mathrm{Cov}(AX, BY) = E\Big[A\,(X - E[X])\,(Y - E[Y])^{\top}\,B^{\top}\Big].$$

Since A and $B^{\top}$ are not random, they can be factored out of the expectation:

$$\mathrm{Cov}(AX, BY) = A\;E\big[(X - E[X])\,(Y - E[Y])^{\top}\big]B^{\top}.$$

By definition:

$$\mathrm{Cov}(X, Y) \;=\; E\Big[(X - E[X])\,(Y - E[Y])^{\top}\Big].$$

Hence,

$$\mathrm{Cov}(AX, BY) = A\;\underbrace{E\big[(X - E[X])\,(Y - E[Y])^{\top}\big]}_{=\,\mathrm{Cov}(X,Y)}\;B^{\top},$$

$$\mathrm{Cov}(AX,\; BY) \;=\; A\,\mathrm{Cov}(X, Y)\,B^{\top}.$$

Question 4

(a)

$$\Pr\big(Z^2 \le t\big) \;=\; \Pr\big(-\sqrt{t} \;\le\; Z \;\le\; \sqrt{t}\big).$$

$$\Pr\big(-\sqrt{t} \le Z \le \sqrt{t}\big) \;=\; \Pr\big(Z \le \sqrt{t}\big) \;-\; \Pr\big(Z < -\sqrt{t}\big).$$

By the symmetry of the standard normal distribution, $\Pr\big(Z < -a\big) = 1 - \Pr\big(Z \le a\big)$. Hence

$$\Pr\big(Z \le \sqrt{t}\big) \;-\; \big[1 - \Pr\big(Z \le \sqrt{t}\big)\big] = 2\,\Pr\big(Z \le \sqrt{t}\big) \;-\; 1.$$

Therefore,

$$\Pr\big(Z^2 \le t\big) = 2\,\Pr\big(Z \le \sqrt{t}\big) \;-\; 1.$$

(b)

From part (a), for $t\, > 0$,

$$F_{Z^2}(t) \;=\; \Pr\big(Z^2 \le t\big) \;=\; 2\,\Pr\big(Z \le \sqrt{t}\big) \;-\; 1.$$

Denote $\Phi(x)$ as the CDF of $N(0,1)$, and $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ as its PDF. Thus

$$F_{Z^2}(t) = 2\,\Phi\big(\sqrt{t}\big) \;-\; 1.$$

Differentiate with respect to t:

$$f_{Z^2}(t) \;=\; \frac{d}{dt}\Big[\,2\,\Phi(\sqrt{t}) - 1\Big] \;=\; 2\,\phi\big(\sqrt{t}\big)\cdot\frac{d}{dt}\big(\sqrt{t}\big).$$

Since $\frac{d}{dt}(\sqrt{t}) = \frac{1}{2\sqrt{t}}$,

$$f_{Z^2}(t) = 2\,\phi\big(\sqrt{t}\big)\,\frac{1}{2\sqrt{t}} = \frac{\phi(\sqrt{t})}{\sqrt{t}}, \quad t > 0.$$

Substitute $\phi(\sqrt{t}) = \frac{1}{\sqrt{2\pi}}\,\exp\big(-\,t/2\big)$:

$$f_{Z^2}(t) = \frac{1}{\sqrt{2\pi\,t}}\,\exp\Big(-\tfrac{t}{2}\Big), \quad t > 0.$$

For $t \le 0$, since $Z^2 \ge 0$, and Z^2 is a continuous random variable:

$$\Pr\big(Z^2 \le t\big) = 0$$

$$f_{Z^2}(t) = 0$$

Hence we confirm that Z^2 has the χ^2_1 distribution with PDF

$$f_{\chi^2_1}(t) = \begin{cases} \dfrac{1}{\sqrt{2\pi\,t}}\;e^{-t/2}, & t > 0, \\ 0, & t \le 0. \end{cases}$$

This matches the known pdf for a chi-square distribution with 1 degree of freedom.

Question 5

(a)

Since:

$$P = \frac{1_n\,1_n^\top}{n},$$

and

$$I_n \text{ is the } n \times n \text{ identity matrix}$$

To prove: $P^2 = P$.

$$P^2 \;=\; \frac{1_n\,1_n^\top}{n}\,\frac{1_n\,1_n^\top}{n}$$

$$P^2 \;=\; \frac{1_n\left(1_n^\top 1_n\right)1_n^\top}{n^2}$$

$$P^2 \;=\; \frac{1_n\,(n)\,1_n^\top}{n^2}$$

Since:

$$1_n^\top\,1_n \;=\; \underbrace{1+1+\cdots+1}_{n \text{ times}} \;=\; n.$$

Therefore,

$$P^2 \;=\; \frac{1_n\,(n)\,1_n^\top}{n^2} \;=\; \frac{1_n\,1_n^\top}{n} \;=\; P.$$

For $(I_n - P)^2 = (I_n - P)$: Since $P^2 = P$ and $I_n^2 = I_n$ by definition:

$$(I_n - P)^2 = I_n^2 - 2P + P^2 \;= I_n - 2P + P = \; I_n - P$$

For $(I_n - P)\,P = 0$:

$$(I_n - P)\,P \;=\; P - P^2 \;=\; P - P \;=\; 0.$$

And, take the transpose:

$$P^\top \;=\; \left(\tfrac{1_n\,1_n^\top}{n}\right)^\top \;=\; \frac{1_n^{\top\top}\,1_n^{\top\top}}{n} \;=\; \frac{1_n\,1_n^\top}{n} \;=\; P.$$

by the fact that $(AB)^\top = B^\top A^\top$ and that $1_n^{\top\top} = 1_n$.

Hence P is symmetric.

Therefore,

$$P^2 = P \quad \text{and} \quad P^\top = P.$$

P is an orthogonal projector onto the subspace spanned by 1_n, and $(I_n - P)$ is the orthogonal projector onto the orthogonal complement.

(b)

Let:

$$\overline{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$$

where $\overline{Z}$ is the sample mean of $Z_1, Z_2, \ldots, Z_n$.

By definition of P,

$$PZ = \frac{1_n \, 1_n^\top}{n} \, Z.$$

$$PZ = \frac{1_n}{n} \big(1_n^\top Z\big),$$

Since $1_n^\top Z = \sum_{i=1}^n Z_i$:

$$PZ = \frac{1}{n} \, 1_n \left(\sum_{i=1}^n Z_i \right) = \overline{Z} \, 1_n.$$

Thus, PZ is a vector where every component equals the sample mean $\overline{Z}$:

$$PZ = \begin{pmatrix} \overline{Z} \\ \overline{Z} \\ \vdots \\ \overline{Z} \end{pmatrix}.$$

By definition:

$$(I_n - P)Z = Z - PZ.$$

Substituting $PZ = \overline{Z} \, 1_n$:

$$(I_n - P)Z = Z - \overline{Z} \, 1_n.$$

$$(I_n - P)Z = \begin{pmatrix} Z_1 - \overline{Z} \\ Z_2 - \overline{Z} \\ \vdots \\ Z_n - \overline{Z} \end{pmatrix}.$$

By the definition:

$$\|x\|^2 = \sum_{i=1}^n x_i^2 = x^\top x.$$

The squared Euclidean norm of $(I_n - P)Z$ is:

$$\|(I_n - P)Z\|^2 = \sum_{i=1}^n \big(Z_i - \overline{Z}\big)^2.$$

For $\overline{Z}$:

Since $PZ = \overline{Z} \, 1_n$:

$$\frac{1}{n} 1_n^\top PZ = \frac{1}{n} 1_n^\top (\overline{Z} \, 1_n).$$

Since $1_n^\top 1_n = n$:

$$\frac{1}{n} 1_n^\top (\overline{Z} \, 1_n) = \frac{1}{n} (1_n^\top 1_n) \overline{Z} = \frac{1}{n} \cdot n \cdot \overline{Z} = \overline{Z}.$$

We have proved:

$$\overline{Z} = \frac{1}{n} 1_n^\top PZ$$

(c)

To prove that $\|(I_n - P)Z\|^2$ (the squared norm of deviations) and $\frac{1}{n}1_n^\top PZ$ (the sample mean) are independent, it suffices to show that the random vectors $(I_n - P)Z$ and PZ are independent, because:

If $(I_n - P)Z$ and PZ are independent, then any function of $(I_n - P)Z$ depends only on $(I_n - P)Z$, and any function of PZ depends only on PZ. Thus, $\|(I_n - P)Z\|^2$ and $\frac{1}{n}1_n^\top PZ$ are independent as they are functions of independent random vectors. Therefore, showing $\mathrm{Cov}((I_n - P)Z, PZ) = 0$ is sufficient to conclude independence.

Let $A = (I_n - P)$ and $B = P$. By the formula proved in Question 3, for covariance of linear transformations:

$$\mathrm{Cov}(AZ, BZ) = A\,\mathrm{Cov}(Z, Z)\,B^\top.$$

The covariance between $(I_n - P)Z$ and PZ is given by:

$$\mathrm{Cov}\big((I_n - P)Z, PZ\big) = (I_n - P)\,\mathrm{Cov}(Z, Z)\,P^\top.$$

Since $\mathrm{Cov}(Z, Z) = Var(Z) = I_n$, we have:

$$\mathrm{Cov}\big((I_n - P)Z, PZ\big) = (I_n - P)I_n P^\top.$$

From part (a), we know that $P^\top = P$, so:

$$(I_n - P)I_n P^\top = (I_n - P)I_n P = (I_n - P)P.$$

Since from part (a), it has been proved that $(I_n - P)P = 0$, it follows that:

$$\mathrm{Cov}\big((I_n - P)Z, PZ\big) = 0.$$

Since $(I_n - P)Z$ and PZ are uncorrelated and $Z \sim N(0, I_n)$, they are independent.