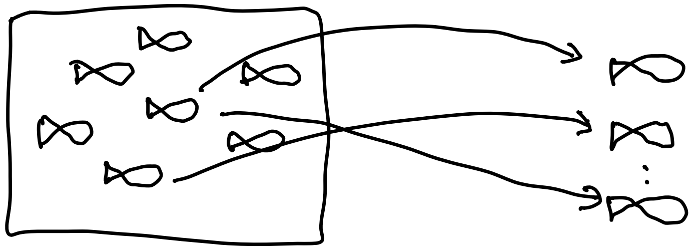
TMA4267 L6

27.01.2017

Parameterestimation: μ and Σ

[H3.3, H4.5]

$\underset{p\times 1}{\underline{X}}$ random vector ← quality of salmon



$\underline{X}_1$ $p\times 1$

$\underline{X}_2$

$\vdots$

$\underline{X}_n$

$p\times 1$ ↓ $\underline{X}_j$ jth obs

population $f(x)$

random sample = independently drawn from the population

3 variations on notation:

a) $\underline{X}_j$, $j=1,\dots,n$

b) $\underset{n\times p}{\underline{X}} = \begin{bmatrix} \underline{X}_1^T \\ \underline{X}_2^T \\ \vdots \\ \underline{X}_n^T \end{bmatrix}$ data matrix rows = obs, cols = RV's

c) $\underset{np\times 1}{\underline{X}^*} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \\ \vdots \\ \underline{X}_n \end{bmatrix}$ vector of RV's

How can we use $\underline{X}_j$, $j=1,\dots,n$ to estimate $\underline{\mu = E(\underline{X})}$ and $\underline{\Sigma = Cov(\underline{X})}$?

① Estimating $\mu = E(\underline{X})$

$$\underset{p\times 1}{\bar{\underline{X}}} = \frac{1}{n}\sum_{j=1}^{n} \underline{X}_j$$

$$\underset{p\times 1}{E(\bar{\underline{X}})} = E\left(\frac{1}{n}\sum_{j=1}^{n}\underline{X}_j\right) = \frac{1}{n}\sum_{j=1}^{n}\overbrace{E(\underline{X}_j)}^{\mu} = \frac{1}{n}\cdot n\cdot\mu = \underline{\mu}$$

unbiased

$$\underset{p\times p}{Cov(\bar{\underline{X}})} = Cov\left(\frac{1}{n}\sum_{j=1}^{n}\underline{X}_j\right)$$

$$Cov(\underline{X}) = E\left((\underline{X}-\mu)(\underline{X}-\mu)^T\right)$$

$$= \left(\frac{1}{n}\right)^2\sum_{j=1}^{n} Cov(\underline{X}_j)$$

$\underline{X}_1, .., \underline{X}_n$ independent

$$= \frac{1}{n^2}\sum_{j=1}^{n}\underbrace{\Sigma}_{Cov(\underline{X}_j)} = \frac{1}{n^2}\cdot n\cdot\Sigma = \underline{\underline{\frac{1}{n}\Sigma}}$$

If $\underline{X}_1, \underline{X}_2, .., \underline{X}_n$ i. i. d $N_p(\mu, \Sigma)$

identically

independent distributed

$$\bar{\underline{X}} \sim N_p\left(\mu, \frac{1}{n}\Sigma\right)$$

② Estimating Σ

Motivation: $\Sigma = E\left((\underline{X}-\mu)(\underline{X}-\mu)^T \right)$

$$\underset{p\times p}{S} = \frac{1}{n-1} \sum_{j=1}^{n} \underbrace{(\underline{X}_j - \bar{\underline{X}})(\underline{X}_j - \bar{\underline{X}})^T}_{p\times p}$$

$$E(S) = \underset{\substack{\uparrow \\ \text{see proof, separate file}}}{\cdots} = \Sigma \quad \text{unbiased}$$

Quadratic forms [F: B.3.3, Theorem B.2]

$\underset{p\times 1}{X}$ random vector, $\underset{p\times p}{A}$ constant matrix

$$\underbrace{\underset{1\times p}{X^T}\,\underset{p\times p}{A}\,\underset{p\times 1}{X}}_{1\times 1} = \sum_{i=1}^{p}\sum_{j=1}^{p} a_{ij}\, x_i \cdot x_j$$ is a quadratic form.

$\mu = E(X)$, $\Sigma = Cov(X)$

The "trace formula":

$$E(X^T A X) = tr(A \underset{\uparrow\, Cov(X)}{\Sigma}) + \mu^T A \mu$$

Ex: V2014 P1a

$p = 3$, $\mu = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\Sigma = I$, $A = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$

$$E(X^T A X) = tr(\underbrace{A I}_{A}) + (1\ 1\ 1)\underbrace{A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{0}$$

$$= 3 \cdot \frac{2}{3} + (1\ 1\ 1)\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \underline{\underline{2}}$$

Symmetric and idempotent matrices

A is symmetric ; $A = A^T$

A is idempotent ; $AA = A$

* The eigenvalues of A are 0 and 1

$$Ae = \lambda e$$
$$AAe = A\lambda e$$
$$Ae = \lambda \underbrace{Ae}_{\lambda e} = \lambda^2 e$$

only 0 and 1 possible

* The rank of a matrix is the number of linearly independent rows, and also given as the number of nonzero eigenvalues.

* General rule: $\underset{n \times n}{tr(A)} = \sum_{i=1}^{n} \lambda_i$, since $\lambda_i = 0$ or 1

tr(A) must be the number of nonzero eigenvalues.

$\Rightarrow$ $tr(A) = rank(A)$ for symmetric and idempotent A.

Now: example = the centering matrix.
(RecEx1. P4, CompEx1. 2a)

$$\underset{n\times n}{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T = \begin{bmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & & \\ \vdots & & \ddots & \\ -\frac{1}{n} & & & 1-\frac{1}{n} \end{bmatrix}$$

$$\begin{bmatrix}1\\1\\1\\1\end{bmatrix} \quad [1\ 1\ 1\ 1]$$

Distribution of quadratic form

$\underset{p\times 1}{X} \sim N_p(0, I)$

$\nwarrow X_1, \ldots, X_p$ are independent

R is symmetric and idempotent with $\text{rank}(R) = r$.

Result: $\quad X^T R X \sim \chi^2_r$

Proof via $\underline{R = P \Lambda P^T}$

$$X^T R X = \text{sum of } r\ N(0,1)^2 \sim \chi^2_r$$

In Comp Ex 1.12c: $X \sim N_n(\mu \mathbf{1}, \sigma^2 I)$

$\Downarrow$

$N_n(0, I)$

Finally: Ratio of quadratic forms
(more in Part 2-3)

---

* $\underline{X} \sim N_p(0, I)$
* R and S symmetric and idempotent with
  $\text{rank}(R) = r$, $\text{rank}(S) = s$
* $RS = 0$

Result: a) $\underline{X}^T R \underline{X} \sim \chi^2_r$

$\underline{X}^T S \underline{X} \sim \chi^2_s$

and independent.

Because: $R\underline{X}$ and $S\underline{X}$ are indep if $R\,\text{Cov}(\underline{X}) S^T = 0$

We have $\text{Cov}(\underline{X}) = I$ and $S = S^T$ so

$R I S^T = RS = 0 \Rightarrow$ yes!

Then a function of $R\underline{X}$ $f(R\underline{X})$ and function of $S\underline{X}$ $g(S\underline{X})$ are also independent $\Rightarrow$ $\underline{X}^T R \underline{X}$ will be indep of $\underline{X}^T S \underline{X}$

(*) see next page for a comment

b)
$$\frac{\frac{\underline{X}^T R \underline{X}}{r}}{\frac{\underline{X}^T S \underline{X}}{s}} \qquad \frac{\frac{\chi^2_r}{r}}{\frac{\chi^2_s}{s}} \sim F_{r,s}$$

This result will be used a lot in Part 2+3.

$\circledast$ Here $f(R\mathbf{X}) = \frac{1}{r}(R\mathbf{X})^T R\mathbf{X} = \frac{1}{r}\mathbf{X}^T \underbrace{R^T R}_{\substack{= RR \quad \text{symmetric } R \\ = R \quad \text{idempotent } R}} \mathbf{X}$

$$= \frac{1}{r}\mathbf{X}^T R \mathbf{X}$$

and the same for $g(S\mathbf{X})$.