L20 Population, sample and statistics

1. Parameters of population, sample and basic statistics

   (1) A univariate case

   Population: $(\mu,\, \sigma^2)$ Sample $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim (\mu 1_n,\, \sigma^2 I_n)$.

   Sample mean: An unbiased estimator for μ

   $$\overline{X} = \frac{\sum X_i}{n} = \frac{1_n'}{n} X \sim \left(\frac{1_n'}{n}\mu 1_n,\, \frac{1_n'}{n}\sigma^2 I_n \frac{1_n}{n}\right) = \left(\mu,\, \frac{\sigma^2}{n}\right).$$

   Sample variance: An unbiased estimator for σ^2:

   $$s^2 = \frac{\sum (X_i - \overline{X})^2}{n-1} = \frac{\sum X_i^2 - \frac{1}{n}(\sum_i X_i)^2}{n-1} = X' \frac{I_n - \frac{1_n 1_n'}{n}}{n-1} X \text{ where } \frac{1_n 1_n'}{n} = 1_n 1_n^+$$

   has mean $(\mu 1_n)' \frac{I_n - 1_n 1_n^+}{n-1}(\mu 1_n) + \operatorname{tr}\left(\frac{I_n - 1_n 1_n^+}{n-1}\sigma^2 I_n\right) = 0 + \sigma^2 = \sigma^2$.

   (2) A multivariate case

   p-dimensional population: $(\mu,\, \Sigma)$ Sample: $X = (X_1, ..., X_n) \sim (\mu 1_n',\, \Sigma,\, I_n)$.

   Sample mean: An unbiased estimator for μ

   $$\overline{X} = \frac{\sum X_i}{n} = X \frac{1_n}{n} \sim \left(\mu 1_n' \frac{1_n}{n},\, \Sigma,\, \frac{1_n'}{n} I_n \frac{1_n}{n}\right) = \left(\mu,\, \Sigma,\, \frac{1}{n}\right) \sim \left(\mu,\, \frac{\sigma^2}{n}\right).$$

   SSCP (Sum of squares and cross products) matrix: $XX' = (\sum_k X_{ik} X_{jk})_{p\times p}$.

   CSSCP (Corrected sum of squares and cross products) matrix

   $(X - \overline{X}1_n')(X - \overline{X}1_n')' = \sum_i (X_i - \overline{X})(X_i - \overline{X})' = \sum_i X_i X_i' - \frac{(\sum X_i)(\sum X_i)'}{n} = X\left(I_n - 1_n 1_n^+\right) X'$

   has mean $(\mu 1_n')\left(I_n - 1_n 1_n^+\right)(\mu 1_n') + \operatorname{tr}\left(I_n - 1_n 1_n^+\right)\Sigma = (n-1)\Sigma$.

   Sample variance-covariance matrix $S = \frac{\text{CSSCP}}{n-1}$ is an unbiased estimator for Σ.

2. Distributions of population sample and basic statistics

   (1) A univariate normal population

   Population: $N(\mu,\, \sigma^2)$ Sample $X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim N(\mu 1_n,\, \sigma^2 I_n)$.

   $\overline{X} = \frac{1_n'}{n} X \sim N\left(\mu,\, \frac{\sigma^2}{n}\right)$ and $s^2 = X' \frac{I_n - 1_n 1_n^+}{n-1} X$ are independent.

   Pf: $\left(\frac{1_n'}{n}\right)(\sigma^2 I_n)\left(\frac{I_n - \frac{1_n 1_n'}{n}}{n-1}\right) = 0$.

   Comment: If $X \sim N(\mu,\, \Sigma)$ and $A\Sigma B = 0$ where $B' = B$, AX and $X'BX$ indep.

   (2) A multivariate normal population

   p-dimensional population: $N(\mu,\, \Sigma)$ Sample: $X = (X_1, ..., X_n) \sim N_{p\times n}(\mu 1_n',\, \Sigma,\, I_n)$.

   Sample mean $\overline{X} = X\frac{1_n}{n} \sim N\left(\mu,\, \frac{\sigma^2}{n}\right)$ and CSSCP $= X\left(I_n - 1_n 1_n^+\right) X'$ are independent.

   Pf: $\left(\frac{1_n}{n}\right)' I_n \left(I_n - 1_n 1_n^+\right) = 0$. So $\overline{X} = X\frac{1_n}{n}$ and $X\left(I_n - 1_n 1_n^+\right)$ are independent.

   Thus $\overline{X}$ and CSSCP $= \left[X\left(I_n - 1_n 1_n^+\right)\right]\left[X\left(I_n - 1_n 1_n^+\right)\right]'$ are independent.

   Comment: Under $X \sim N_{p\times n}(M,\, \Sigma,\, \Psi)$,

   AXB and CXD are independent if and only if $A\Sigma C' = 0$ or $B'\Psi D = 0$.

3. Tools

(1) Parameters

For random vectors $X \sim (\mu,\, \Sigma)$,

$$AX + b \sim (A\mu + b,\, A\Sigma A') \text{ and } E(X'BX) = \mu' B\mu + \mathrm{tr}(B\Sigma).$$

For random vectors X and Y, $E(X'AY) = \mu_x' A\mu_y + \mathrm{tr}(A\Sigma_{YX})$.

For random matrix $X \sim (M,\, \Sigma,\, \Psi)$,

$$AXB \sim (AMB,\, A\Sigma A',\, B'\Psi B)$$
$$E(XAX') = MAM' + \mathrm{tr}(A\Psi)\Sigma$$
$$X' \sim (M',\, \Psi,\, \Sigma)$$
$$X'BX = M'BM + \mathrm{tr}(B\Sigma)\Psi$$

(2) Distributions

With normal vector $X \sim N(\mu,\, \Sigma)$

AX and BX are indep $\Longleftrightarrow A\Sigma B' = 0$

$A\Sigma B = 0$ where $B' = B \Longrightarrow AX$ and $X'BX$ are indep

$A\Sigma B = 0$ where $A' = A$ and $B' = B \Longrightarrow X'AX$ and $X'BX$ are indep

With normal matrix $X \sim N(M,\, \Sigma,\, \Psi)$

AXB and CXD are indep $\Longleftrightarrow A\Sigma C' = 0$ or $B'\Psi D = 0$.

L21 Chi-square and Wishart distributions

1. A sampling distribution: χ^2-distribution

   (1) Definition
   If $Z \sim N(\mu,\, I_k)$, the distribution of $Z'Z = Z_1^2 + \cdots + Z_k^2$ is called a χ^2-distribution with non-centrality parameter $\mu'\mu$ and degrees of freedom k denoted by $\chi^2(\mu'\mu,\, k)$.

   $$Z \sim N(\mu,\, I_k) \Longrightarrow Z'Z \sim \chi^2(\mu'\mu,\, k).$$

   The notation for central χ^2-distribution $\chi^2(0,\, k)$ is simplified as $\chi^2(k)$.

   (2) A theorem
   If $X \sim N(\mu,\, \Sigma)$ where $\Sigma > 0$ and $A\Sigma A = A = A'$, then $X'AX \sim \chi^2(\mu' A\mu,\, \mathrm{tr}(A\Sigma))$.

   Pf: $A\Sigma A = A = A' \Longrightarrow \Sigma^{1/2}A\Sigma^{1/2}$ is symmetric and idempotent.
   In the compact form of EVD $\quad \Sigma^{1/2}A\Sigma^{1/2} = P_r P_r', \quad P_r'P_r = I_r$ and $r = \mathrm{tr}(A\Sigma)$.

   $$\begin{aligned} X \sim N(\mu,\, \Sigma) &\Longrightarrow Z = P_r'\Sigma^{-1/2}X \sim N(P_r'\Sigma^{-1/2}\mu,\, I_r) \\ &\Longrightarrow Z'Z \sim \chi^2((P_r'\Sigma^{-1/2}\mu)'(P_r'\Sigma^{-1/2}\mu),\, r) = \chi^2(\mu'A\mu,\, \mathrm{tr}(A\Sigma)). \\ X'AX &= X'\Sigma^{-1/2}(\Sigma^{1/2}A\Sigma^{1/2})\Sigma^{-1/2}X = X'\Sigma^{-1/2}P_rP_r'\Sigma^{-1/2}X \\ &= (P_r'\Sigma^{-1/2}X)(P_r'\Sigma^{-1/2}X)' = Z'Z \sim \chi^2(\mu'A\mu,\, \mathrm{tr}(A\Sigma)). \end{aligned}$$

   (3) Distribution associated with sample variance
   s^2 is the sample variance based on a sample of size n from a normal population $N(\mu,\, \sigma^2)$. Then $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$.

   Pf: Note that the sample $X \sim N\left(\mu 1_n,\, \sigma^2 I_n\right)$ and $s^2 = \frac{1}{n-1}X'\left(I_n - 1_n 1_n^+\right)X$.
   So $\frac{(n-1)s^2}{\sigma^2} = X'AX$ where $A = \frac{I_n - 1_n 1_n^+}{\sigma^2}$. But

   (i) $A(\sigma^2 I_n)A = A \quad$ (ii) $(\mu 1_n)'A(\mu 1_n) = 0 \quad$ (iii) $\mathrm{tr}(A\sigma^2 I_n) = n-1$
   So $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(0,\, n-1) = \chi^2(n-1)$.

2. A sampling distribution: Wishart distribution

   (1) Definition
   If $Z \sim N_{p\times n}(M,\, \Sigma,\, I_n)$, the distribution of ZZ' is called a Wishart distribution with non-centrality parameter MM', degrees of freedom n and a parameter matrix Σ denoted by $W_{p\times p}(MM',\, n,\, \Sigma)$.

   $$Z \sim N_{p\times n}(M,\, \Sigma,\, I_n) \Longrightarrow ZZ' \sim W_{p\times p}(MM',\, n,\, \Sigma).$$

   The notation for central Wishart $W_{p\times p}(0,\, n,\, \Sigma)$ is simplified as $W_{p\times p}(n,\, \Sigma)$. The notation for standard Wishart $W_{p\times p}(n,\, I_p)$ is simplified as $W_{p\times p}(n)$.

   (2) A theorem (Proof is left for you)
   If $X \sim N_{p\times n}(M,\, \Sigma,\, I_n)$ and $A^2 = A = A' \in R^{n\times n}$, $XAX' \sim W_{p\times p}(MAM',\, \mathrm{tr}(A),\, \Sigma)$.

   (3) Distribution associated with sample variance-covariance matrix
   S is the sample variance-covariance matrix based on a sample of size n from a normal population $N(\mu,\, \Sigma)$. Then $(n-1)S \sim W_{p\times p}(n-1,\, \Sigma)$.

Pf: Note that the sample $X \sim N_{p\times n}\,(M,\,\Sigma,\,I_n)$ where $M = \mu 1_n'$ and
$(n-1)S = \text{CSSCP} = XAX'$ where $A = I_n - 1_n 1_n^+$. But
(i) $A^2 = A = A'$ (so that XAX' has Wishart distribution)
(ii) $MAM' = 0$ (This is the value for the non-centrality parameter)
(iii) $\text{tr}(A) = n-1$ (This is the degrees of freedom).
So $(n-1)S = \text{CSSCP} \sim W_{p\times p}(n-1,\,\Sigma)$. Conclusion follows.