L09: Sampling distributions of SSCP, CSSCP, S and S_u

1. Wishart distributions

    (1) Wishart distribution

    Recall: $X' \sim N_{p \times n}(M, \Sigma, I_n) \implies X'X \sim W_{p \times p}(MM', \Sigma, n)$.
    Distribution relation: $[N_{p \times n}(M, \Sigma, I_n)][N_{p \times n}(M, \Sigma, I_n)]' = W_{p \times p}(MM', \Sigma, n)$.
    When determining the parameters in Wishart we follow that in the first normal matrix factor.

    Ex1: Using $Y \sim N_{m \times n}(M, \Sigma, \Psi)$ to construct a matrix with Wishart distribution.
        $Y \sim N_{m \times n}(M, \Sigma, \Psi) \implies Y\Psi^{-1/2} \sim N_{m \times n}(M\Psi^{-1/2}, \Sigma, I_n)$.
        So $(Y\Psi^{-1/2})(Y\Psi^{-1/2})' \sim W_{m \times m}\left((M\psi^{-1/2})(M\Psi^{-1/2})', \Sigma, n\right)$, i.e.,

    $$Y\Psi^{-1}Y' \sim W_{m \times m}(M\Psi^{-1}M', \Sigma, n).$$

    (2) Relation to χ^2-distribution
    $\chi^2(\nu, k) = W_{1 \times 1}(\nu, 1, k)$

    Proof Recall: $\mathbf{x} \in N(\mu, I_k) \implies \mathbf{x}'\mathbf{x} \sim \chi^2(\mu'\mu, k)$. So

    $$
    \begin{aligned}
    \chi^2(\mu'\mu, k) &= [N(\mu, I_k)]'[N(\mu, I_k)] = [N_{k \times 1}(\mu, I_k, 1)]'[N_{k \times 1}(\mu, I_k, 1)] \\
    &= [N_{1 \times k}(\mu', 1, I_k)][N_{1 \times k}(\mu', 1, I_k)]' \sim W_{1 \times 1}(\mu'\mu, 1, k).
    \end{aligned}
    $$

    So $\chi^2(\nu, k) = W_{1 \times 1}(\nu, 1, k)$.
    Ex2: $\chi^2(k) = \chi^2(0, k) = W_{1 \times 1}(0, 1, k) = W_{1 \times 1}(k)$.

    (3) Lemma
    $X' \sim N_{p \times n}(M, \Sigma, I_n)$ and $A' = A^2 = A \in R^{n \times n}$ with $\text{rank}(A) = r$.
    Then $X'AX \sim W_{p \times p}(MAM', \Sigma, r)$.

    Proof If $A' = A^2 = A \in R^{n \times n}$ with $\text{rank}(A) = \text{tr}(A) = r$,
        then $A = PP'$ where $P \in R^{n \times r}$ and $P'P = I_r$. So $X'P \sim N_{p \times r}(MP, \Sigma, I_r)$.
        Thus $X'AX = (X'P)(X'P)' \sim W_{p \times p}((MP)(MP)', \Sigma, r) = W_{p \times p}(MAM', \Sigma, r)$.

    (4) Theorem
    $X' \sim N_{m \times n}(M, \Sigma, \Psi)$. $A' = A \in R^{n \times n}$, $A\Psi A = A$ and $\text{rank}(A) = r$.
    Then $X'AX \sim W_{m \times m}(MAM', \Sigma, r)$.

    Proof $X' \sim N_{m \times n}(M, \Sigma, \Psi) \implies X'\Psi^{-1/2} \sim N_{m \times n}(M\Psi^{-1/2}, \Sigma, I_n)$.
        $A' = A \implies (\Psi^{1/2}A\Psi^{1/2})' = \Psi^{1/2}A\Psi^{1/2}$; $A\Psi A = A \implies (\Psi^{1/2}A\Psi^{1/2})^2 = \Psi^{1/2}A\Psi^{1/2}$
        and $\text{rank}(\Psi^{1/2}A\Psi^{1/2}) = \text{rank}(A) = r$. By the last Lemma
        $(X'\Psi^{-1/2})(\Psi^{1/2}A\Psi^{1/2})(X'\Psi^{-1/2})' \sim W_{m \times m}\left((M\Psi^{-1/2})(\Psi^{1/2}A\Psi^{1/2})(M\Psi^{-1/2}), \Sigma, r\right)$.
        Thus $X'AX \sim W_{m \times m}(MAM', \Sigma, r)$.
    Ex3: Back to Ex1, $Y \sim N_{m \times n}(M, \Sigma, \Psi)$. $(\Psi^{-1})' = \Psi^{-1}$, $\Psi^{-1}\Psi\Psi^{-1} = \Psi^{-1}$ and $\text{rank}(\Psi^{-1}) = n$.
        Thus $Y'\Psi^{-1}Y \sim W_{m \times m}(M\Psi^{-1}M', \Sigma, r)$. This verifies the result in Ex1.

2. Expectation of Wishart distributions

    (1) $E(x'Ay) = \mu_x'A\mu_y + \text{tr}(A\Sigma_{yx})$.
    $E(\mathbf{x}) = \mu_x \in R^p$, $E(\mathbf{y}) = \mu_y \in R^q$ and $\text{Cov}(\mathbf{y}, \mathbf{x}) = \Sigma yx \in R^{q \times p}$.
    Then $E(\mathbf{x}'A\mathbf{y}) = \mu_x'A\mu_y + \text{tr}(A\Sigma_{yx})$.

    Proof Note that $\Sigma_{yx} = \text{Cov}(y, x) = E(yx') - \mu_y\mu_x'$. So

    $$
    \begin{aligned}
    E(x'Ay) &= E[\text{tr}(x'Ay)] = E[\text{tr}(yx'A)] = \text{tr}[E(yx'A)] = \text{tr}[E(yx')A] \\
    &= \text{tr}[(\Sigma_{yx} + \mu_y\mu_x')A] = \text{tr}(\mu_y\mu_x'A) + \text{tr}(\Sigma_{yx}A) = \mu_x'A\mu_y + \text{tr}(A\Sigma_{yx}).
    \end{aligned}
    $$

Ex4: If $x \sim (\mu, \Sigma)$, $E(x'Ax) = \mu'A\mu + \text{tr}(A\Sigma)$. This is a formula we saw before.

(2) Theorem

If $X \in R^{m \times n}$ and $X \sim (M, \Sigma, \Psi)$, then $E(X'AX) = M'AM + \text{tr}(A\Sigma)\Psi$ and
$$E(XBX') = MBM' + \text{tr}(B\Psi)\Sigma.$$

Proof Sketch:
$$E(X'AX) = E[(e'_{in}X'AXe_{jn})_{n \times n}] = \cdots = M'AM + \text{tr}(A\Sigma)\Psi$$
Comment: Determine $\text{tr}(A\Sigma)\Psi$ and $\text{tr}(B\Psi)\Sigma$ by the dimensions of $X'AX$ and XBX'.
Ex5: From $x \sim (\mu, \Sigma) = (\mu, \Sigma, 1)$, $E(x'Ax) = \mu'A\mu + \text{tr}(A\Sigma) \cdot 1$.
$$E(xbx') = \mu b\mu' + \text{tr}(b \cdot 1)\Sigma \overset{b \neq 0}{\Longrightarrow} E(xx') = \mu\mu' + \Sigma \Longrightarrow \Sigma = E(xx') - \mu\mu'.$$

(3) $E[W_{p \times p}(D, \Sigma, k)] = D + k\,\Sigma$

Proof $W \sim W_{p \times p}(D, \Sigma, k) \Longrightarrow W = X'X$ where $X' \sim N_{p \times k}(M, \Sigma, I_k)$ and $MM' = D$.
$E(W) = E(X'I_kX) = MM' + \text{tr}(I_k \cdot I_k)\Sigma = D + k\,\Sigma$.
Ex6: $E[W_{p \times p}(\Sigma, k)] = K\,\Sigma$ and $E[W_{p \times p}(k)] = E[W_{p \times p}(I_p, k)] = kI_p$.

3. Sampling distributions

$X \in R^{n \times p}$ is data matrix from $N(\mu, \Sigma)$.
So $X' = (\mathbf{x}_1, .., \mathbf{x}_n) \sim N_{p \times n}(\mu 1'_n, \Sigma, I_n)$.

(1) Distribution of SSCP

$\text{SSCP} = \sum_{i=1}^{n} \mathbf{x}_i\mathbf{x}'_i = X'X \sim W_{p \times p}(n\mu\mu', \Sigma, n)$. $E(SSCP) = n(\mu\mu' + \Sigma)$.

(2) Distribution of CSSCP

$\text{CSSCP} = \sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})' = X'\left(I_n - \frac{1_n 1'_n}{n}\right)X$.

$I_n - \frac{1_n 1'_n}{n}$ is symmetric idempotent with rank $n-1$.
So $CSSCP \sim W_{p \times p}(0, \Sigma, n-1) = W_{p \times p}(\Sigma, n-1)$. $E(CSSCP) = (n-1)\Sigma$.

(3) Theorem

$W \sim W_{p \times p}(D, \Sigma, k)$ and $A \in R^{q \times p} \Longrightarrow AWA' \sim W_{q \times q}(ADA', A\Sigma A', k)$.

Proof $W \sim W_{p \times p}(D, \Sigma, k) \Longrightarrow W = X'X$, $X' \sim N_{p \times k}(M, \Sigma, I_k)$, $D = MM'$.
So $AX' \sim N_{q \times k}(AM, A\Sigma A', I_k)$ and
$AWA' = (AX')(AX')' \sim W_{p \times p}(AM(AM)', A\Sigma A', k) = W_{q \times q}(ADA', A\Sigma A', k)$.
Ex7: With $CSSCP \sim W_{p \times p}(\Sigma, n-1)$,
$S = \frac{CSSCP}{n} = \frac{I_p}{\sqrt{n}}CSSCP\frac{I_p}{\sqrt{n}} \sim W_{p \times p}\left(\frac{\Sigma}{n}, n-1\right)$. $E(S) = \frac{n-1}{n}\Sigma$.
$S_u = \frac{CSSCP}{n-1} = \frac{I_p}{\sqrt{n-1}}CSSCP\frac{I_p}{\sqrt{n-1}} \sim W_{p \times p}\left(\frac{\Sigma}{n-1}, n-1\right)$. $E(S_u) = \Sigma$

L10 Hotelling's T^2-distribution

1. Hotelling's T^2-distributions

   (1) Definitions

   Suppose $X_0 \sim N(\mu,\, I_p) = N_{p \times 1}(\mu,\, I_p,\, 1)$ is independent to $W \sim W_{p \times p}(k)$.
   Then $T^2 = X_0' \left(\frac{W}{k}\right)^{-1} X_0 \sim T^2(\mu,\, p,\, k)$ is called a Hotelling's T^2-distribution with
   non-centrality parameter vector μ from X_0, dimension parameter p from both X_0 and W, and
   degrees of freedom k from W.

   Comment: $T^2(0,\, p,\, k) = T^2(p,\, k)$ is called a central T^2-distribution

   Ex1: If $X_0 \sim N(\mu,\, I_p)$ is independent to $W \sim W_{p \times p}(k)$, then
   $(X_0 - \mu)' \left(\frac{W}{k}\right)^{-1} (X_0 - \mu) \sim T^2(p,\, k)$.

   (2) Relations
   $T^2(\mu,\, 1,\, k) = [t(\mu,\, k)]^2 = F(\mu^2,\, 1,\, k)$.

   Proof If $T^2 \sim T^2(\mu,\, 1,\, k)$, then $T^2 = X_0 \left(\frac{W}{k}\right)^{-1} X_0$ where
   $\qquad\qquad X_0 \sim N(\mu,\, 1^2)$ is independent to $W \sim W_{1 \times 1}(k) = \chi^2(k)$.
   But $T^2 = X_0' \left(\frac{W}{k}\right)^{-1} X_0 = \left(\frac{X_0}{\sqrt{W/k}}\right)^2$ and $T^2 = \frac{X_0^2/1}{W/k}$ where

   $\qquad\qquad \frac{X_0}{\sqrt{W/k}} \sim t(\mu,\, k)$ and $\frac{X_0^2/1}{W/k} \sim F(\mu^2,\, 1,\, k)$.
   Thus $T^2(\mu,\, 1,\, k) = [t(\mu,\, k)]^2 = F(\mu^2,\, 1,\, k)$.

   Comment: $T^2(1,\, k) = [t(k)]^2 = F(1,\, k)$.

   (3) Theorem
   Suppose $X_0 \sim N_p(\mu,\, \Sigma)$ is independent to $W \sim W_{p \times p}(\Sigma,\, k)$. Then

   $$X_0' \left(\frac{W}{k}\right)^{-1} X_0 \sim T^2(\Sigma^{-1/2}\mu,\, p,\, k).$$

   Proof Note that $X_0' \left(\frac{W}{k}\right)^{-1} X_0 = (\Sigma^{-1/2} X_0)' \left(\frac{\Sigma^{-1/2} W \Sigma^{-1/2}}{k}\right)^{-1} (\Sigma^{-1/2} X_0)$.
   But $X_0 \sim N_p(\mu,\, \Sigma) \Longrightarrow \Sigma^{-1/2} X_0 \sim N(\Sigma^{-1/2}\mu,\, I_p)$;
   $\qquad W \sim W_{p \times p}(\Sigma,\, k) \Longrightarrow \Sigma^{-1/2} W \Sigma^{-1/2} \sim W_{p \times p}(I_p,\, k) = W_{p \times p}(k)$;
   $\qquad X_0$ and W are independent $\Longrightarrow \Sigma^{-1/2} X_0$ and $\Sigma^{-1/2} W \Sigma^{-1/2}$ are independent.
   Thus $X_0' \left(\frac{W}{k}\right)^{-1} X_0 = (\Sigma^{-1/2} X_0)' \left(\frac{\Sigma^{-1/2} W \Sigma^{-1/2}}{k}\right)^{-1} (\Sigma^{-1/2} X_0) \sim T^2(\Sigma^{-1/2}\mu,\, p,\, k)$

   Corollary: Suppose $X_0 \sim N_p(\mu,\, \Sigma)$ is independent to $W \sim W_{p \times p}(\Sigma,\, k)$. Then

   $$(X_0 - \mu)' \left(\frac{W}{k}\right)^{-1} (X_0 - \mu) \sim T^2(p,\, k).$$

2. Distribution of $T^2 = (\overline{\mathbf{x}} - \mu_0)' \left(\frac{S_u}{n}\right)^{-1} (\overline{\mathbf{x}} - \mu_0)$.

   (1) Theorem
   $X \sim N_{m \times n}(M,\, \Sigma,\, \Psi)$.

   $$AXB \text{ and } CXD \text{ are independent} \iff A\Sigma C' = 0 \text{ or } B'\Psi D = 0$$

   Proof Skipped.

   Ex2: $\overline{\mathbf{x}}$ and CSSCP of sample from $N(\mu,\, \Sigma)$ are independent.
   $\qquad$ Let $X \in R^{n \times p}$ be data matrix for a sample from $N(\mu,\, \Sigma)$. Then $X' \sim N_{p \times n}(\mu 1_n',\, \Sigma,\, I_n)$.

So $X'\left(\frac{1_n}{n}\right)$ and $X'\left(I_n - \frac{1_n 1'_n}{n}\right)$ are independent since $\left(\frac{1_n}{n}\right)' I_n \left(I_n - \frac{1_n 1'_n}{n}\right) = 0$.

Hence $X'\left(\frac{1_n}{n}\right)$ and $\left[X'\left(I_n - \frac{1_n 1'_n}{n}\right)\right]\left[X'\left(I_n - \frac{1_n 1'_n}{n}\right)\right]'$ are independent.

Thus $\bar{\mathbf{x}}$ and CSSCP are independent.

(2) $(\bar{\mathbf{x}} - \mu_0)' \left(\frac{S_u}{n}\right)^{-1} (\bar{\mathbf{x}} - \mu_0) \sim T^2(\Sigma^{-1/2}\sqrt{n}(\mu - \mu_0),\, p,\, n-1)$.

Proof Let $\bar{\mathbf{x}}$ and CSSCP be from $N(\mu,\, \Sigma)$. Then
$\bar{\mathbf{x}} \sim N\left(\mu,\, \frac{\Sigma}{n}\right) \Longrightarrow \sqrt{n}(\bar{\mathbf{x}} - \mu_0) \sim N(\sqrt{n}(\mu - \mu_0),\, \Sigma)$;
CSSCP $\sim W_{p \times p}(\Sigma,\, n-1)$; $\bar{\mathbf{x}}$ and CSSCP are independent.
So $\sqrt{n}(\bar{\mathbf{x}} - \mu_0)' \left(\frac{CSSCP}{n-1}\right)^{-1} \sqrt{n}(\bar{\mathbf{x}} - \mu_0) \sim T^2(\Sigma^{-1/2}\sqrt{n}(\mu - \mu_0),\, p,\, n-1)$.
Thus $(\bar{\mathbf{x}} - \mu) \left(\frac{S_u}{n}\right)^{-1} (\bar{\mathbf{x}} - \mu_0) \sim T^2(\Sigma^{-1/2}\sqrt{n}(\mu - \mu_0),\, p,\, n-1)$.

Ex3: With $\mu_0 = \mu$, $(\bar{\mathbf{x}} - \mu)' \left(\frac{S_u}{n}\right)^{-1} (\bar{\mathbf{x}} - \mu) \sim T^2(p,\, n-1)$.

3. Relations to F-distributions

(1) Theorem
$$T^2(\mu,\, p,\, k) = \frac{pk}{k-p+1} F(\mu'\mu,\, p,\, k-p+1)$$
$$T^2(p,\, k) = \frac{pk}{k-p+1} F(p,\, k-p+1)$$

Proof Skipped.

Comment: There are tables for central F-distributions in most textbooks, but not for non-central F-distributions.

Ex4: $T^2(\mu,\, 1,\, k) = \frac{1 \times k}{k-1+1} F(\mu'\mu,\, 1,\, k-1+1) = F(\mu^2,\, 1,\, k)$. Result in (2) of 1 is verified.

(2) Based on a sample from $N(\mu,\, \Sigma)$,
$$\begin{aligned}(\bar{\mathbf{x}} - \mu_0)' \left(\frac{S_u}{n}\right)^{-1} (\bar{\mathbf{x}} - \mu_0) \;&\sim\; T^2(\Sigma^{-1/2}\sqrt{n}(\mu - \mu_0),\, p,\, n-1) \\ &=\; \frac{(n-1)p}{n-p} F(n(\mu - \mu_0)'\Sigma^{-1}(\mu - \mu_0),\, p,\, n-p).\end{aligned}$$
$$(\bar{\mathbf{x}} - \mu)' \left(\frac{S_u}{n}\right)^{-1} (\bar{\mathbf{x}} - \mu) \sim T^2(p,\, n-1) = \frac{(n-1)p}{n-p} F(p,\, n-p).$$

Ex5: Let $T^2 = (\overline{X} - \mu)' \left(\frac{S_u}{n}\right)^{-1} (\overline{X} - \mu)$ with $p = 3$ and $n = 10$. Find $P(T^2 > 4)$.

$$\begin{aligned}T^2 \;&=\; (\overline{X} - \mu)' \left(\frac{S_u}{n}\right)^{-1} (\overline{X} - \mu) \\ &\sim\; T^2(p,\, n-1) = \frac{p(n-1)}{n-p} F(p,\, n-p) = \frac{3 \times 9}{10-3} F(3,\, 10-3) = \frac{27}{7} F(3,\, 7).\end{aligned}$$
$$P(T^2 > 4) = P\left(\frac{27}{7} F(3,\, 7) > 4\right) = P\left(F(3,\, 7) > \frac{28}{27}\right) = P(F(3,\, 7) > 1.037) = 0.4331$$

Ex6: $W \sim W_{1 \times 1}(25, 5)$. Find $P(W > 10)$.
$W \sim W_{1 \times 1}(25,\, 5) \Longrightarrow \frac{1}{5} W \frac{1}{5} \sim \frac{1}{5} W_{1 \times 1}(25,\, 5) \frac{1}{5} = W_{1 \times 1}(1,\, 5) = W_{1 \times 1}(5) = \chi^2(5)$.
So $\frac{W}{25} \sim \chi^2(5)$. Thus $P(W > 10) = P\left(\chi^2(5) > \frac{10}{25}\right) = P(\chi^2(5) > 0.4) = 0.99533$.