L11: One-sample estimation

1. Point estimators

   (1) Likelihood function
   $\mathbf{x}_i$, $i = 1, .., n$ is a random sample from $N(\mu, \Sigma)$. From this sample

$$\begin{aligned} L(\mu, \Sigma) &= \prod_{i=1}^{n} \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left[-\tfrac{1}{2}(\mathbf{x}_i - \mu)\Sigma^{-1}(\mathbf{x}_i - \mu)\right] \\ &= \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left\{-\tfrac{1}{2}\sum_{i=1}^{n}[(\mathbf{x}_i - \overline{x} + \overline{x} - \mu)\Sigma^{-1}(\mathbf{x}_i - \overline{x} + \overline{x} - \mu)\right] \\ &= \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left[-\tfrac{1}{2}(\overline{x} - \mu)'\left(\tfrac{\Sigma}{n}\right)^{-1}(\overline{x} - \mu)\right] \exp\left[-\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1/2}\,\mathrm{CSSCP}\,\Sigma^{-1/2})\right]. \end{aligned}$$

   (2) Maximum likelihood estimator (MLE) for μ

   Clearly $L(\mu, \Sigma) \leq L(\overline{x}, \Sigma) = \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left[-\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1/2}\,\mathrm{CSSCP}\,\Sigma^{-1/2})\right]$ for all μ and Σ.
   So $\overline{x}$ is MLE for μ.

   (3) MLE for Σ

   By EVD, $\Sigma^{-1/2}\,\mathrm{CSSCP}\,\Sigma^{-1/2} = P\Lambda P'$.

$$\begin{aligned} L(\overline{x}, \Sigma) &= \frac{|\Sigma^{-1/2}\,\mathrm{CSSCP}\,\Sigma^{-1/2}|^{n/2}}{(2\pi)^{np/2}|\mathrm{CSSCP}|^{n/2}} \exp\left[-\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1/2}\,\mathrm{CSSCP}\,\Sigma^{-1/2})\right] \\ &= \frac{(\lambda_1 \cdots \lambda_p)^{n/2}}{(2\pi)^{np/2}|\mathrm{CSSCP}|^{n/2}} \exp\left(-\frac{\lambda_1 + \cdots + \lambda_p}{2}\right) = \frac{1}{(2\pi)^{np/2}|\mathrm{CSSCP}|^{n/2}} \prod_{i=1}^{p} g(\lambda_i). \end{aligned}$$

   $g(\lambda_i) = \lambda_i^{n/2} e^{-\frac{\lambda_i}{2}}$, $\ln g(\lambda_i) = \frac{n}{2}\ln\lambda_i - \frac{\lambda_i}{2}$, $[\ln g(\lambda_i)]' = \frac{n}{2}\frac{1}{\lambda_i} - \frac{1}{2} \overset{set}{=\!=} 0 \Longrightarrow \lambda_i = n$.
   $[\ln g(\lambda_i)]'' = \frac{n}{2}\frac{-1}{\lambda_i^2} < 0$. So $g(\lambda_i)$ is maximized at $\lambda_i = n$. Hence $L(\overline{x}, \Sigma)$ is maximized at

$$\lambda_i = n \text{ for all } i \Longleftrightarrow \Sigma^{-1/2}\,\mathrm{CSSCP}\,\Sigma^{-1/2} = nI \Longleftrightarrow \Sigma = \frac{\mathrm{CSSCP}}{n} = S.$$

   So $S = \frac{\mathrm{CSSCP}}{n}$ is MLE for Σ.
   Comment: $L(\overline{x}, S) = \left(\frac{n}{2\pi e}\right)^{np/2} |\mathrm{CSSCP}|^{-n/2}$.

   (4) Properties

   $\overline{x} \sim N\left(\mu, \frac{\Sigma}{n}\right) \Longrightarrow E(\overline{x}) = \mu \Longrightarrow \overline{x}$ is an unbiased estimator (UE) for μ.
   $S \sim W_{p\times p}\left(\frac{\Sigma}{n}, n-1\right) \Longrightarrow E(S) = \frac{n-1}{n}\Sigma \neq \Sigma \Longrightarrow S$ is a biased estimator for Σ.
   $S_u \sim W_{p\times p}\left(\frac{\Sigma}{n-1}, n-1\right) \Longrightarrow E(S_u) = \Sigma \Longrightarrow S_u$ is an UE for Σ.

2. Confidence intervals

   (1) $\frac{l'\overline{x} - l'\mu}{\sigma_{l'\overline{x}}} \sim N(0, 1^2)$

   $\overline{x} \sim N\left(\mu, \frac{\Sigma}{n}\right) \Longrightarrow l'\overline{x} \sim N(l'\mu, \sigma^2_{l'\overline{x}})$ where $\sigma^2_{l'\overline{x}} = \frac{l'\Sigma l}{n}$. So $\frac{l'\overline{x} - l'\mu}{\sigma_{l'\overline{x}}} \sim N(0, 1^2)$.

   (2) $\frac{(l'S_u l)(n-1)}{l'\Sigma l} \sim \chi^2(n-1)$.

   $S_u \sim W_{p\times p}\left(\frac{\Sigma}{n-1}, n-1\right) \Longrightarrow l'S_u l \sim W_{1\times 1}\left(\frac{l'\Sigma l}{n-1}, n-1\right) \Longrightarrow \frac{(l'S_u l)(n-1)}{l'\Sigma l} \sim W_{1\times 1}(n-1)$, i.e.,
   $\frac{(l'S_u l)(n-1)}{l'\Sigma l} \sim \chi^2(n-1)$.

(3) $\frac{l'\overline{x}-l'\mu}{s_{l'\overline{x}}} \sim t(n-1)$.

$\sigma^2_{l'\overline{x}} = \frac{l'\Sigma l}{n}$ has UE $s^2_{l'\overline{x}} = \frac{l'S_u l}{n}$. $s_{l'\overline{x}} = \sqrt{\frac{l'S_u l}{n}}$ is estimated standard deviation of $l'\overline{x}$.

$\overline{x}$ and S_u are independent$\Longrightarrow \frac{l'\overline{x}-l'\mu}{\sigma_{l'\overline{x}}\sqrt{\frac{l'S_u l}{l'\Sigma l}}} \sim t(n-1)$.

But $\sigma_{l'\overline{x}}\sqrt{\frac{l'S_u l}{l'\Sigma l}} = \sqrt{\frac{l'S_u l}{n}} = s_{l'\overline{x}}$. Thus $\frac{l'\overline{x}-l'\mu}{s_{l'\overline{x}}} \sim t(n-1)$.

(4) $1-\alpha$ confidence interval for $l'\mu$

$l'\overline{x} \pm t_{\alpha/2}(n-1)s_{l'\overline{x}}$ is a $1-\alpha$ C.I. for $l'\mu$.

Proof
$$\begin{aligned} 1-\alpha &= P(-t_{\alpha/2}(n-1) < t(n-1) < t_{\alpha/2}(n-1)) \\ &= P\left(-t_{\alpha/2}(n-1) < \tfrac{l'\overline{x}-l'\mu}{s_{l'\overline{x}}} < t_{\alpha/2}(n-1)\right) \\ &= P\left(l'\overline{x} - t_{\alpha/2}(n-1)s_{l'\overline{x}} < l'\mu < l'\overline{x} + t_{\alpha/2}(n-1)s_{l'\overline{x}}\right). \end{aligned}$$

Ex1: With $l = e_i$, $e_i'\mu = \mu_i$; $e_i'\overline{x} = \overline{x}_i$, $s^2_{\overline{x}_i} = \frac{(S_u)_{ii}}{n}$.

$$\overline{x}_i \pm t_{\alpha/2}(n-1)s_{\overline{x}_i} \text{ is a } 1-\alpha \text{ confidence interval for } \mu_i.$$

(5) Simultaneous CIs: Bonferroni method

Suppose there are k confidence intervals, each has confidence coefficient $1-\frac{\alpha}{k}$. Then these k intervals are simultaneously true with probability $\geq 1-\alpha$.

Proof Let I_i be the ith confidence interval. Then $P(I_i) \geq 1-\frac{\alpha}{k}$, $i = 1, ..., k$. So
$$\begin{aligned} P(I_1 \cap \cdots \cap I_k) &= 1 - P((I_1 \cap \cdots \cap I_k)^c) = 1 - P(I_1^c \cup \cdots \cup I_k^c) \\ &\geq 1 - [P(I_1^c) + \cdots + P(I_k^c)] = 1 - [(1-P(I_1)) + \cdots + (1-P(I_k))] \\ &= 1 - k + [P(I_1) + \cdots + P(I_k)] \geq 1 - k + k\left(1-\tfrac{\alpha}{k}\right) = 1-\alpha. \end{aligned}$$

Ex2: $l_i'\overline{x} \pm t_{\alpha/(2k)}(n-1)s_{l_i'\overline{x}}$, $i = 1, .., k$, are simultaneous confidence intervals for $l_i'\mu$, $i = 1, .., k$, with overall confidence coefficient $1-\alpha$.

3. Confidence region for μ

$\left\{\mu \in R^p : (\overline{x}-\mu)'\left(\frac{S_u}{n}\right)^{-1}(\overline{x}-\mu) \leq T^2_\alpha(p, n-1)\right\}$
is a confidence region for μ with confidence coefficient $1-\alpha$.

Proof $P\left((\overline{x}-\mu)'\left(\frac{S_u}{n}\right)^{-1}(\overline{x}-\mu) \leq T^2_\alpha(p, n-1)\right) = 1-\alpha$.

Comments: This region is an ellipsoid in R^p with center $\overline{x}$.
$T^2_\alpha(p, n-1) = \frac{(n-1)p}{n-p}F_\alpha(p, n-p)$.

L12: Schefee's intervals

1. Recall
   Consider $\mu \in R^p$ in $N(\mu, \Sigma)$.

   (1) t-intervals

   $l'\overline{x} \pm t_{\alpha/2}(n-1)s_{l'\overline{x}}$ is a $1-\alpha$ CI for $l'\mu$.
   Here $s_{l'\overline{x}} = \sqrt{\frac{l'S_u l}{n}}$ is estimated standard deviation of $l'\overline{x}$.
   $\overline{x}_i \pm t_{\alpha/2}(n-1)s_{\overline{x}_i}$ is a $1-\alpha$ CI for μ_i.
   Here $s_{\overline{x}_i} = \sqrt{\frac{(S_u)_{ii}}{n}}$ is estimated standard deviation of $\overline{x}_i$.

   (2) Bonferroni intervals

   $l_i'\overline{x} \pm t_{\alpha/(2k)}(n-1)s_{l_i'\overline{x}}$, $i = 1, .., k$, are simultaneous CIs for $l_i'\mu$, $i = 1, .., k$, with overall confidence coefficient $1-\alpha$.
   $\overline{x}_i \pm t_{\alpha/(2k)}(n-1)s_{\overline{x}_i}$, $i = 1, .., k$, are simultaneous CIs for μ_i, $i = 1, .., k$, with overall confidence coefficient $1-\alpha$.

   (3) Confidence region

   $\left\{\mu \in R^p : (\mu - \overline{x})' \left(\frac{S_u}{n}\right)^{-1} (\mu - \overline{x}) \le T^2_\alpha(p, n-1)\right\}$
   is a confidence region for μ with confidence coefficient $1-\alpha$.

2. One-sided confidence intervals

   (1) Lower-sided confidence interval

   $(-\infty, l'\overline{x} + t_\alpha(n-1)s_{l'\overline{x}})$ is a lower-sided $1-\alpha$ confidence interval for $l'\mu$.
   Proof Note that $\frac{l'\overline{x}-l'\mu}{s_{l'\overline{x}}} \sim t(n-1) \Longleftrightarrow \frac{l'\mu-l'\overline{x}}{s_{l'\overline{x}}} \sim t(n-1)$.

   $$\begin{aligned} 1-\alpha &= P(-\infty < t(n-1) < t_\alpha(n-1)) = P\left(-\infty < \frac{l'\mu-l'\overline{x}}{s_{l'\overline{x}}} < t_\alpha(n-1)\right) \\ &= P(-\infty < l'\mu < l'\overline{x} + t_\alpha(n-1)s_{l'\overline{x}}). \end{aligned}$$

   Ex1: $(-\infty, \overline{x}_i + t_\alpha(n-1)s_{\overline{x}_i})$ is a $1-\alpha$ CI for μ_i.

   (2) Upper-sided confidence interval

   $(l'\overline{x} - t_\alpha(n-1)s_{l'\overline{x}}, \infty)$ is an upper-sided $1-\alpha$ confidence interval for $l'\mu$.
   Proof Note that $\frac{l'\mu-l'\overline{x}}{s_{l'\overline{x}}} \sim t(n-1)$.

   $$\begin{aligned} 1-\alpha &= P(-t_\alpha(n-1) < t(n-1) < \infty) = P\left(-t_\alpha(n-1) < \frac{l'\mu-l'\overline{x}}{s_{l'\overline{x}}} < \infty\right) \\ &= P(l'\overline{x} - t_\alpha(n-1)s_{l'\overline{x}} < l'\mu < \infty). \end{aligned}$$

   Ex2: $(\overline{x}_i - t_\alpha(n-1)s_{\overline{x}_i}, \infty)$ is a $1-\alpha$ CI for μ_i.

3. Schefee's simultaneous CIs

   (1) Extended Cauchy-Schwartz inequality
   For $x, y \in R^p$, $(x'y)^2 \le (x'x)(y'y)$ is Cauchy-Schwartz inequality.
   With $A > 0$, for $A^{-1/2}x$ and $A^{1/2}x$,
   $(x'y)^2 \le (x'A^{-1}x)(y'Ay)$ is extended Cauchy-Schwartz inequality.

(2) A lemma

If $x'A^{-1}x \le c$, then $-\sqrt{c\,(y'Ay)} \le x'y \le \sqrt{c\,(y'Ay)}$.

Proof Suppose $x'A^{-1}x \le c$. By extended Cauchy-Schwartz inequality,

$$(x'y)^2 \le (x'A^{-1}x)(y'Ay) \le c\,(y'Ay).$$

So $-\sqrt{c\,(y'Ay)} \le x'y \le \sqrt{c,\,(y'Ay)}$.

(3) Schefee's simultaneous CIS

$l_i'\overline{x} \pm \sqrt{T_\alpha^2(p,\,n-1)}\,s_{l_i'\overline{x}}$, $i = 1, 2, ...,$
are simultaneous CIs for $l_i'\mu$, $i = 1, 2, ...,$ with overall confidence coefficient $1-\alpha$.

Proof In Lemma with $x = \overline{x} - \mu$, $y = l_i$, $A = \frac{S_u}{n}$ and $c = T_\alpha^2(p, n-1)$,

$$\begin{aligned}
\text{Let } E_0 &= [x'A^{-1}x \le c] = \left[(\overline{x}-\mu)'\left(\tfrac{S_u}{n}\right)^{-1}(\overline{x}-\mu) \le T_\alpha^2(p,\,n-1)\right].\\
\text{Let } E_i &= [-\sqrt{c,\,(y'Ay)} \le x'y \le \sqrt{c\,(y'Ay)}\ ]\\
&= \left[-\sqrt{T_\alpha^2(p,\,n-1)l_i'\tfrac{S_u}{n}l_i} \le l_i'\mu - l_i'\overline{x} \le \sqrt{T_\alpha^2(p,\,n-1)\,l_i'\tfrac{S_u}{n}l_i}\right]\\
&= \left[-\sqrt{T_\alpha^2(p,n-1)}\,s_{l_i'\overline{x}} \le l_i'\mu - l_i'\overline{x} \le \sqrt{T_\alpha^2(p,\,n-1)}\,s_{l_i'\overline{x}}\right]\\
&= \left[l_i'\overline{x} - \sqrt{T_\alpha^2(p,n-1)}\,s_{l_i'\overline{x}} \le l_i'\mu \le l_i'\overline{x} + \sqrt{T_\alpha^2(p,\,n-1)}\,s_{l_i'\overline{x}}\right].
\end{aligned}$$

By the Lemma, $E_0 \subset E_i \Longrightarrow E_0 \subset \cap_i E_i \Longrightarrow P(E_0) \le P(E_1 \cap E_2 \cap \cdots)$.
But $P(E_0) = 1-\alpha$. So $P(E_1 \cap E_2 \cap \cdots) = 1-\alpha$.

Comments: Consider CI for $l'\mu$.
If the CI is for $l'\mu$ only with confidence coefficient $1-\alpha$, then the width of the interval is $w_1 = 2t_{\alpha/2}(n-1)s_{l'\overline{x}}$.
If the CI is one of k Bonferroni intervals with overall confidence coefficient $1-\alpha$, then the width of the interval is $w_2 = 2t_{\alpha/(2k)}(n-1)s_{l'\overline{x}}$.
If the CI is one of Schefee's interval with overall confidence coefficient $1-\alpha$, then the width of the CI is $w_3 = 2\sqrt{T_\alpha^2(p,n-1)}\,s_{l'\overline{x}}$.

$T_\alpha^2(p,n-1) = \frac{(n-1)p}{n-p}F_\alpha(p,n-p)$.