5.4.1 Direct method to identify complete sufficient statistics

$\boxed{\textit{Example 5.4.1.}}$. Let $X_1, \dots, X_n$ be a random sample from $\mathcal{P}oi(\theta)$, $\Theta = \{\theta : \theta > 0\}$. Find UMVUE using sufficiency and completeness.

Follow the summary of constructing UMVUE:

Step 1: (Method 1)

Step 1.1: Joint distribution is

$$f(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} \left(\frac{\theta^{x_i} e^{-\theta}}{x_i!} \mathbb{1}_{\{0,1,\dots\}}(x_i) \right) = \underbrace{\frac{\prod_{i=1}^{n} \mathbb{1}_{\{0,1,\dots\}}(x_i)}{\prod_{i=1}^{n} x_i!}}_{=h(x_1,\dots,x_n)} \underbrace{\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta}}_{=g(\sum_{i=1}^{n} x_i;\theta)}$$

By factorization criterion, $S = \sum_{i=1}^{n} X_i$ is sufficient.

Step 1.2: We know that $S \sim \mathcal{P}oi(n\theta)$, so $f_S(s;\theta) = \frac{(n\theta)^s e^{-n\theta}}{s!}$, $s = 0, 1, \dots$. Want to check that $\{f_S(s;\theta); \theta \in \Theta\}$ is a complete. Take U and assume $\mathbb{E}(U(S)) = 0 \ \forall \theta$

$$\implies \sum_{s=0}^{\infty} U(s) \frac{(n\theta)^s e^{-n\theta}}{s!} = 0, \quad \forall \theta \in \Theta$$

$$\implies \sum_{s=0}^{\infty} \frac{U(s) n^s}{s!} \theta^s = 0, \quad \forall \theta \in \Theta$$

By uniqueness of power series expansion

$$\implies \frac{U(s) n^s}{s!} = 0 \quad \text{for } s = 0, 1, \dots$$

$$\implies U(s) = 0 \quad \text{for } s = 0, 1, \dots$$

$$\implies U(S) = 0 \quad \text{with probability 1} \quad \forall \theta \in \Theta$$

So S is complete.

Step 2: (Method 1) $\mathbb{E}(S/n) = \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) = \theta$.

Step 3: By Lehmann-Scheffe theorem, S/n is UMVUE for θ.

$\boxed{\text{Example 5.4.2.}}$. Let $X_1, \ldots, X_n$ be a random sample from $\mathcal{U}nif(0,\theta)$, $\Theta = \{\theta : \theta > 0\}$. Find UMVUE for θ using sufficiency and completeness.

Step 1: (Method 1)

Step 1.1: We have shown in Example 5.1.3 that $\max\limits_{1\le i\le n} X_i$ is sufficient statistic for θ.

Step 1.2: To check for completeness, we need the density of S

$$\mathbb{P}\{S \le s\} = \mathbb{P}\{\max_{1\le i\le n} X_i \le s\} = \begin{cases} \left(\frac{s}{\theta}\right)^n, & 0 \le s \le \theta \\ 1, & s > \theta \\ 0, & s < 0 \end{cases}$$

So $f_S(s;\theta) = \frac{n}{\theta^n} s^{n-1} \mathbb{1}_{[0,\theta]}(s)$. Let U be a function such that $\mathbb{E}U(S) = 0\ \forall \theta \in \Theta$.

$$\begin{aligned}
&\implies \int_0^\theta U(s) \frac{n}{\theta^n} s^{n-1}\, ds = 0 \quad \forall \theta \in \Theta \\
&\implies \int_0^\theta U(s) n s^{n-1}\, ds = 0 \quad \forall \theta \in \Theta \\
&\implies U(\theta) n\theta^{n-1} = 0 \quad \forall \theta \in \Theta \quad \text{by differentiation} \\
&\implies U(\theta) = 0 \quad \forall \theta \in \Theta = \{\theta : \theta > 0\} \\
&\implies U(s) = 0 \quad \text{for } 0 \le s \le \theta \\
&\implies U(S) = 0 \quad \text{with probability 1}
\end{aligned}$$

So S is complete sufficient.

Step 2: (Method 1) Since $\mathbb{E}S = \int_0^\theta s \frac{n}{\theta^n} s^{n-1}\, ds = \frac{n}{\theta^n} \frac{\theta^{n+1}}{n+1} = \frac{n\theta}{n+1}$, $T^* = \frac{n+1}{n} S$ is unbiased for θ.

Step 3: By Lehmann-Scheffe theorem, $T^* = \frac{n+1}{n} S$ is UMVUE for θ.

Note that CRLB is not applicable here, because $\{x : f(x;\theta) = 0\}$ depends on θ.

$\boxed{\textit{Example 5.4.3.}}$. Let $X_1, \ldots, X_n$ be a random sample from $\mathcal{DU}(\theta)$, i.e.

$$\mathbb{P}\{X_1 = x\} = \frac{1}{\theta}, \; x = 1, 2, \ldots, \theta$$

θ is unknown, $\Theta = \{\theta : \theta = 1, 2, \ldots\}$ Find UMVUE for θ using sufficiency and completeness.

Step 1: (Method 1)

Step 1.1:

$$f(x;\theta) = \frac{1}{\theta}\mathbb{1}_{\{1,2,\ldots,\theta\}}(x)$$

$$f(x_1, \ldots, x_n; \theta) = \prod_{i=1}^{n}\left(\frac{1}{\theta}\mathbb{1}_{\{1,2,\ldots,\theta\}}(x_i)\right) = \underbrace{\frac{1}{\theta^n}\mathbb{1}_{\{1,2,\ldots,\theta\}}(\max_{1\le i\le n} x_i)}_{=g(\max_{1\le i\le n} x_i;\theta)} \underbrace{\prod_{i=1}^{n}\mathbb{1}_{\{1,2,\ldots\}}(x_i)}_{=h(x_1,\ldots,x_n)}$$

So $S = \max_{1\le i\le n} X_i$ is sufficient statistic for θ.

Step 1.2: To check for completeness, we need the distribution of S

$$\begin{aligned}
\mathbb{P}\{S = s\} &= \mathbb{P}\{S \le s\} - \mathbb{P}\{S \le s-1\} \\
&= \mathbb{P}\{X_1 \le s, \ldots, X_n \le s\} - \mathbb{P}\{X_1 \le s-1, \ldots, X_n \le s-1\} \\
&= \left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n = f_S(s;\theta), \quad s \in \{1, 2, \ldots, \theta\}
\end{aligned}$$

Take U and assume $\mathbb{E}(U(S)) = 0 \; \forall\theta \in \Theta$, which implies

$$\mathbb{E}(U(S)) = \sum_{s=1}^{\theta} U(s)\left(\left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n\right) = 0, \quad \forall\theta \in \Theta$$

Take $\theta = 1$

$$\sum_{s=1}^{1} U(s)\left(s^n - (s-1)^n\right) = 0 \implies U(1) = 0$$

Take $\theta = 2$

$$\sum_{s=1}^{2} U(s)\left(\left(\frac{s}{2}\right)^n - \left(\frac{s-1}{2}\right)^n\right) = \underbrace{U(1)}_{=0}\left(\frac{1}{2}\right)^n + U(2)\left(1 - \left(\frac{1}{2}\right)^n\right) = 0$$

$$\implies U(2) = 0$$

Continue on this, we will get $U(1) = U(2) = \ldots = 0$, so $U(S) = 0$ with probability 1 $\forall \theta \in \Theta = \{\theta : \theta = 1, 2, \ldots\}$. So S is complete.

Step 2: (Method 2) Now because $\mathbb{E}X_1 = \frac{1}{\theta}\sum_{i=1}^{\theta} i = \frac{\theta(\theta+1)}{2\theta} = \frac{\theta+1}{2}$, $2X_1 - 1$ is unbiased for θ.

Step 3: Therefore, by Lehmann-Scheffe theorem, UMVUE is $\mathbb{E}(2X_1 - 1|S)$. One can show this is equal to $\frac{S^{n+1}-(S-1)^{n+1}}{S^n-(S-1)^n}$[1].

5.4.2 Theorem for easily identifying complete sufficient statistics

Definition 5.4.1. **(Regular exponential class).** Let X have PDF or PMF

$$f(x;\theta) = c(\theta)h(x)\exp\left\{\sum_{j=1}^{k} q_j(\theta)t_j(x)\right\}$$

where $\theta = \{\theta_1, \ldots, \theta_k\} \in \Theta = \{\theta : a_i \le \theta_i \le b_i, 1 \le i \le k, -\infty \le a_i < b_i \le \infty\}$ and if it satisfies

1. h depends only on x (not on θ)

[1]Because $\mathbb{P}\{S = s\} = \left(\frac{s}{\theta}\right)^n - \left(\frac{s-1}{\theta}\right)^n$ and $\mathbb{P}\{S = s|X_1 = i\} = \mathbb{P}\{S \le s|X_1 = i\} - \mathbb{P}\{S \le s-1|X_1 = i\} = \begin{cases} \left(\frac{s}{\theta}\right)^{n-1} - \left(\frac{s-1}{\theta}\right)^{n-1}, & i = 1, \ldots, s-1 \\ \mathbb{P}\{S \le s|X_1 = s\} - 0 = \left(\frac{s}{\theta}\right)^{n-1}, & i = s \end{cases}$ and $\mathbb{P}\{X_1 = i\} = \frac{1}{\theta}$.

So $\mathbb{P}\{X_1 = i|S = s\} = \begin{cases} 0, & \text{otherwise} \\ \dfrac{\mathbb{P}\{S = s|X_1 = i\}\mathbb{P}\{X_1 = i\}}{\mathbb{P}\{S = s\}} \end{cases} = \begin{cases} \dfrac{s^{n-1} - (s-1)^{n-1}}{s^n - (s-1)^n}, & i = 1, \ldots, s-1 \\ \dfrac{s^{n-1}}{s^n - (s-1)^n}, & i = s \end{cases}$

So $\mathbb{E}(2X_1 - 1|S = s) = 2\left(\sum_{i=1}^{s} i\mathbb{P}\{X_1 = i|S = s\}\right) - 1$
$= 2\left(\sum_{i=1}^{s-1} i\frac{s^{n-1} - (s-1)^{n-1}}{s^n - (s-1)^n} + s\frac{s^{n-1}}{s^n - (s-1)^n}\right) - 1 = 2\left(\frac{s^{n-1} - (s-1)^{n-1}}{s^n - (s-1)^n}\frac{s(s-1)}{2} + \frac{s^n}{s^n - (s-1)^n}\right) - 1 = \frac{s^{n+1} - (s-1)^{n+1}}{s^n - (s-1)^n}$.

2. range of $(q_1(\theta), \dots, q_k(\theta))$ as θ ranges over Ⓗ contains an open set in $\mathbb{R}^k$

Then X is said to have a distribution in the regular exponential class (REC)

Theorem 5.4.1. **(REC Theorem).** Let $X_1, \dots, X_n$ be a random sample from $f(x;\theta)$ that is in REC. Then $S_1 = \sum_{i=1}^n t_1(X_i), \dots, S_k = \sum_{i=1}^n t_k(X_i)$ are jointly complete sufficient for $(\theta_1, \dots, \theta_k)$.

Proof. Here we proof sufficient only by factorization criterion

$$f(x_1, \dots, x_n; \theta) = \underbrace{h(x_1) \cdots h(x_n)}_{h} \underbrace{c^n(\theta) \exp\left\{ \sum_{j=1}^{k} q_j(\theta) \sum_{i=1}^{n} t_j(x_i) \right\}}_{g}$$

Completeness requires advanced treatment. □

Example 5.4.4. $X_1, \dots, X_n$ are i.i.d. $\mathcal{B}in(1, p)$, Ⓗ $= \{p : 0 < p < 1\}$. Find complete sufficient statistic for p.

$$f(x;p) = \begin{cases} p^x(1-p)^{1-x}, & x \in \{0,1\} \\ 0, & \text{otherwise} \end{cases} = \left(\frac{p}{1-p}\right)^x (1-p) \mathbb{1}_{\{0,1\}}(x)$$

$$= \underbrace{e^{x \ln\left(\frac{p}{1-p}\right)}}_{e^{q_1(p) t_1(x)}} \underbrace{(1-p)}_{c(p)} \underbrace{\mathbb{1}_{\{0,1\}}(x)}_{h(x)}$$

So $t_1(x) = x$ and $q_1(p) = \ln\left(\frac{p}{1-p}\right)$. And $\{q_1(p) : 0 < p < 1\}$ contains an open set in $\mathbb{R}$. By Theorem 5.4.1, $S_1 = \sum_{i=1}^n t_1(X_i) = \sum_{i=1}^n X_i$ is complete sufficient for p.

Example 5.4.5. $X_1, \dots, X_n$ are i.i.d. $\mathcal{N}(\mu, \sigma^2)$, $\theta = (\mu, \sigma^2)$, Ⓗ $= \{\mu, \sigma^2 : \mu \in \mathbb{R}, \sigma^2 > 0\}$. Construct UMVUE.

Step 1: (Method 2)

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x^2 - 2x\mu + \mu^2)}$$

$$= \underbrace{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\mu^2}{2\sigma^2}}}_{c(\mu,\sigma^2)} \underbrace{e^{-\frac{x^2}{2\sigma^2}+\frac{x\mu}{\sigma^2}}}_{e^{q_1(\mu,\sigma^2)t_1(x)+q_2(\mu,\sigma^2)t_2(x)}} \qquad \text{(in this problem, } h(x)=1)$$

So $t_1(x) = x^2$, $t_2(x) = x$, $q_1(\mu,\sigma^2) = \frac{-1}{2\sigma^2}$, $q_2(\mu,\sigma^2) = \frac{\mu}{\sigma^2}$, and $\{q_1(\mu,\sigma^2), q_2(\mu,\sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0\}$ contains an open set in $\mathbb{R}^2$. By the theorem 5.4.1, $S_1 = \sum_{i=1}^n t_1(X_i) = \sum_{i=1}^n X_i^2$, $S_2 = \sum_{i=1}^n t_2(X_i) = \sum_{i=1}^n X_i$ are joint complete sufficient for (μ,σ^2).

Step 2: (Method 1) $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, $S^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2$ are functions of S_1 and S_2 and $\mathbb{E}\overline{X} = \mu$[1], $\mathbb{E}S^2 = \sigma^2$[2].

Step 3: By Lehmann-Scheffe theorem, $\overline{X}$ and S^2 are UMVUE for μ, σ^2.

Note that $\overline{X}$ achieves CRLB of μ[3], σ^2/n for variance since $\mathbb{V}ar\overline{X} = \sigma^2/n$[4]. But $\mathbb{V}arS^2 = \frac{2\sigma^4}{n-1}$[5] $> \frac{2\sigma^4}{n} =$ CRLB of σ^2[6].

---

[1] $\mathbb{E}\overline{X} = \mathbb{E}\frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{n}\sum_{i=1}^n \mathbb{E}X_i = \frac{1}{n}n\mu = \mu$

[2] By Theorem 2.4.2 (iii) $\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$ and mean of chi square, $\mathbb{E}S^2 = \frac{\sigma^2}{n-1}\mathbb{E}(\chi^2(n-1)) = \frac{\sigma^2}{n-1}n-1 = \sigma^2$

[3] $\tau(\mu) = \mu$, $\ln f(X,\mu) = \ln\frac{1}{\sqrt{2\pi}} - \ln\sigma - \frac{1}{2\sigma^2}(X-\mu)^2$, $\frac{\partial}{\partial\mu}\ln f(X,\mu) = \frac{X-\mu}{\sigma^2}$, $\frac{\partial^2}{\partial\mu^2}\ln f(X,\mu) = \frac{-1}{\sigma^2}$, CRLB of $\mu = \frac{1}{-n\mathbb{E}(-1/\sigma^2)} = \frac{\sigma^2}{n}$

[4] $\mathbb{V}ar\overline{X} = \mathbb{V}ar\frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{n^2}\sum_{i=1}^n \mathbb{V}arX_i = \frac{1}{n^2}n\sigma^2 = \sigma^2$

[5] By Theorem 2.4.2 (iii) $\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$ and variance of chi square, $\mathbb{V}arS^2 = \frac{\sigma^4}{(n-1)^2}\mathbb{V}ar(\chi^2(n-1)) = \frac{\sigma^4}{(n-1)^2}2(n-1) = \frac{2\sigma^4}{n-1}$

[6] $\tau(\sigma^2) = \sigma^2$, $\ln f(X,\sigma^2) = \ln\frac{1}{\sqrt{2\pi}} - \ln\sigma - \frac{1}{2\sigma^2}(X-\mu)^2$, $\frac{\partial}{\partial\sigma^2}\ln f(X,\sigma^2) = \frac{-1}{2\sigma^2} + \frac{(X-\mu)^2}{2(\sigma^2)^2}$, $\frac{\partial^2}{\partial(\sigma^2)^2}\ln f(X,\sigma^2) = \frac{1}{2(\sigma^2)^2} - \frac{(X-\mu)^2}{(\sigma^2)^3}$, CRLB of $\sigma^2 = \frac{1}{-n\mathbb{E}\left(\frac{1}{2(\sigma^2)^2} - \frac{(X-\mu)^2}{(\sigma^2)^3}\right)} = \frac{1}{-n\left(\frac{1}{2(\sigma^2)^2} - \frac{\sigma^2}{(\sigma^2)^3}\right)} = \frac{2\sigma^4}{n}$

6

INTERVAL ESTIMATION

6.1 Confidence Intervals

Let $X_1, \dots, X_n$ have joint PDF or PMF $f(x_1, \dots, x_n; \theta)$, $\theta \in \Theta$ = an interval. Let $L = l(X_1, \dots, X_n)$ and $U = u(X_1, \dots, X_n)$ be two statistics. Observed values $X_1 = x_1, \dots, X_n = x_n$ givens observed values $l(x_1, \dots, x_n)$ and $u(x_1, \dots, x_n)$.

Definition 6.1.1. **(Confidence interval)**. An interval $(l(x_1, \dots, x_n), u(x_1, \dots, x_n))$ is called a $100\gamma\%$ confidence interval for θ if the statistic L and U satisfy

$$\mathbb{P}\{l(X_1, \dots, X_n) < \theta < u(X_1, \dots, X_n)\} = \gamma \quad \forall \theta \in \Theta$$

where $\gamma \in (0, 1)$ is fixed. The observed values $l(x_1, \dots, x_n)$ and $u(x_1, \dots, x_n)$ are called lower and upper $100\gamma\%$ confidence limits, respectively.

If

$$\mathbb{P}\{l(X_1, \dots, X_n) < \theta\} = \gamma \quad \forall \theta \in \Theta$$

then $l(x_1, \dots, x_n)$ is called a one-sided lower $100\gamma\%$ confidence limit for θ.

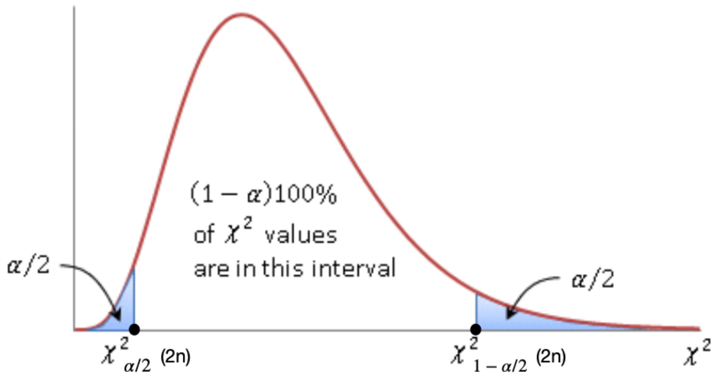
If

$$\mathbb{P}\{\theta < u(X_1, \dots, X_n)\} = \gamma \quad \forall \theta \in \Theta$$

then $u(x_1, \dots, x_n)$ is called a one-sided upper $100\gamma\%$ confidence limit for θ.

Example 6.1.1.. Let $X_1, \dots, X_n$ be a random sample from $\mathcal{E}xp(\theta)$, $\theta \in \Theta = \{\theta : \theta > 0\}$. One can verify that $\frac{2n\overline{X}}{\theta} \sim \chi^2(2n)$[1]. Hence, for all $\theta \in \Theta$.

[1] MGF of X_1 is $\frac{1}{1-\theta t}$, so MGF of $\sum_{i=1}^n X_i$ is $\frac{1}{(1-\theta t)^n}$, so MGF of $\frac{2\sum_{i=1}^n X_i}{\theta} = \frac{2n\overline{X}}{\theta}$ is $\frac{1}{(1-\theta\frac{2t}{\theta})^n} = \frac{1}{(1-2t)^{2n/2}}$ which is $\chi^2(2n)$'s MGF.



$$\mathbb{P}\left\{\chi^2_{\alpha/2}(2n) < \frac{2n\overline{X}}{\theta} < \chi^2_{1-\alpha/2}(2n)\right\} = 1-\alpha \quad \forall\theta \in Ⓗ$$

Now invent the probability statement to get θ in the middle

$$\mathbb{P}\left\{\theta\chi^2_{\alpha/2}(2n) < 2n\overline{X} < \theta\chi^2_{1-\alpha/2}(2n)\right\} = 1-\alpha \quad \forall\theta \in Ⓗ$$

$$\mathbb{P}\left\{\frac{2n\overline{X}}{\chi^2_{1-\alpha/2}(2n)} < \theta < \frac{2n\overline{X}}{\chi^2_{\alpha/2}(2n)}\right\} = 1-\alpha \quad \forall\theta \in Ⓗ$$

So $L = l(X_1, \dots, X_n) = \frac{2n\overline{X}}{\chi^2_{1-\alpha/2}(2n)}$ and $U = u(X_1, \dots, X_n) = \frac{2n\overline{X}}{\chi^2_{\alpha/2}(2n)}$. Hence, if $X_1 = x_1, \dots, X_n = x_n$ are observed, $\left(\frac{2n\overline{x}}{\chi^2_{1-\alpha/2}(2n)}, \frac{2n\overline{x}}{\chi^2_{\alpha/2}(2n)}\right)$ is $100(1-\alpha)\%$ confidence interval for θ.

Now form an numerical example: $x_1 = 2574$, $x_2 = 1310$, $x_3 = 282$, $x_4 = 1233$, $x_5 = 1925$, $x_6 = 135$, $x_7 = 281$, $x_8 = 2254$, $x_9 = 671$, $x_{10} = 495$. So $\overline{x} = 1116$ is UMVU estimate (and MLE) of θ. Take $1-\alpha = 0.80$ or $\alpha/2 = 0.10$, $n = 10$, $\chi^2_{0.1}(20) = 12.44$ and $\chi^2_{0.9}(20) = 28.41$. So 80% confidence interval is $\left(\frac{2\cdot10\cdot1116}{28.41} = 786, \frac{2\cdot10\cdot1116}{12.44} = 1794\right)$.

Therefore, our best point estimate is $\hat{\theta} = \overline{x} = 1116$ and we can be 80% confident that true $\theta \in (786, 1794)$.

6.2 Pivot Method

Definition 6.2.1. **(Pivotal Quantity)**. If $Q = q(X_1, \ldots, X_n; \theta)$ is a random variable depending on $X_1, \ldots, X_n$ and θ such that distribution of Q does not depend on θ, Q is called a pivotal quantity.

Example 6.2.1.. Recall in the last example 6.1.1, $Q = q(X_1, \ldots, X_n; \theta) = \frac{2n\overline{X}}{\theta} \sim \chi^2(2n)$ was pivotal quantity.

Summary 6.2.1. **(Pivotal Method)**. In general, we can form

$$\mathbb{P}\{q_1 < \underbrace{q(X_1, \ldots, X_n; \theta)}_{Q} < q_2\} = \gamma \quad \forall \theta$$

and try to invent the inequality. I.e.

$$\{\theta : q_1 < q(x_1, \ldots, x_n; \theta) < q_2\}$$

forms a $100\gamma\%$ confidence region, and such a confidence region may not be an interval and it might be quite complicated.

If for fixed $x_1, \ldots, x_n$, $q(x_1, \ldots, x_n; \theta)$ is monotonic in θ, then

$$\{\theta : q_1 < q(x_1, \ldots, x_n; \theta) < q_2\}$$

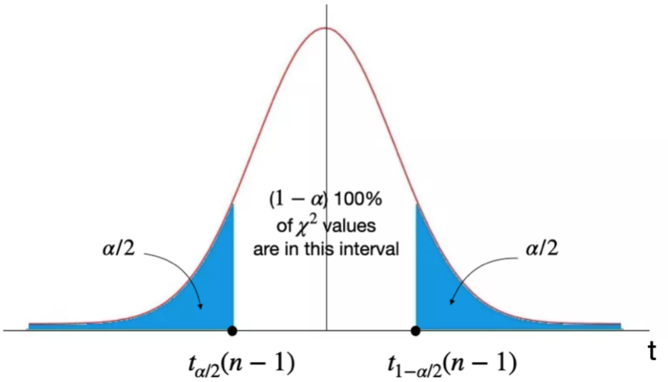
forms an interval, a $100\gamma\%$ confidence interval.

But what are some natural choices for pivotal quantities?

Theorem 6.2.1.. Let $X_1, \ldots, X_n$ be a random sample from $f(x; \theta)$, $\theta \in \Theta$ and assume that MLE of θ, $\hat{\theta}$, exists

1. If $f(x; \theta) = (1/\theta) g(x/\theta)$ for some g, then $\hat{\theta}/\theta$ is pivotal quantitiy.
2. If $f(x; \theta) = g(x - \theta)$ for some g, then $\hat{\theta} - \theta$ is pivotal quantity.

Theorem 6.2.2.. Let $X_1, \ldots, X_n$ be a random sample from $f(x; \theta_1, \theta_2)$, $(\theta_1, \theta_2) \in \Theta = \{\theta_1 \in \mathbb{R}, \theta_2 > 0\}$ where $f(x; \theta_1, \theta_2) = \frac{1}{\theta_2} g\left(\frac{x - \theta_1}{\theta_2}\right)$ for some g, and if MLEs $\hat{\theta}_1$ and $\hat{\theta}_2$ of θ_1, θ_2 exist, then



1. $(\hat{\theta}_1 - \theta_1)/\hat{\theta}_2$ is pivotal quantity for θ_1
2. $\hat{\theta}_2/\theta_2$ is pivotal quantity for θ_2

$\boxed{\textit{Example 6.2.2.}}$. Let $X_1, \ldots, X_n$ be a random sample from $\mathcal{N}(\mu, \sigma^2)$

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Recall HW4, 3), the MLE's are $\hat{\mu} = \frac{1}{n}\sum_{i=1}^n X_i = \overline{X}$, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2 = \frac{n-1}{n}S^2$. By Theorem 6.2.2 ($\theta_2 = \sigma$, $\theta_1 = \mu$, $g(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$), pivotal quantity for μ is $\frac{\overline{X}-\mu}{\hat{\sigma}} = \frac{\overline{X}-\mu}{\sqrt{\frac{n-1}{n}}S}$ or better (take advantage of known distribution, i.e. Corollary 2.5.1) get $\frac{\sqrt{n}(\overline{X}-\mu)}{S} \sim t(n-1)$. Form a confidence interval

$$\mathbb{P}\left\{t_{\alpha/2}(n-1) < \frac{\sqrt{n}(\overline{X}-\mu)}{S} < t_{1-\alpha/2}(n-1)\right\} = 1 - \alpha$$

$$\mathbb{P}\left\{\overline{X} - \frac{t_{1-\alpha/2}(n-1)}{\sqrt{n}}S < \mu < \overline{X} - \frac{t_{\alpha/2}(n-1)}{\sqrt{n}}S\right\} = 1 - \alpha$$

or[1]

$$\mathbb{P}\left\{\overline{X} + \frac{t_{\alpha/2}(n-1)}{\sqrt{n}}S < \mu < \overline{X} + \frac{t_{1-\alpha/2}(n-1)}{\sqrt{n}}S\right\} = 1 - \alpha$$

[1]From the plot, $t_{\alpha/2}(n-1) = -t_{1-\alpha/2}(n-1)$ because student t-distribution is symmetric.

So given observed values $x_1, \ldots, x_n$ and let $\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i$,
then $\left(\overline{x} + \frac{t_{\alpha/2}(n-1)}{\sqrt{n}} s, \overline{x} + \frac{t_{1-\alpha/2}(n-1)}{\sqrt{n}} s\right)$ is $100(1-\alpha)\%$ confidence interval for μ.
Pivotal quantity for σ^2 is $\hat{\sigma}^2/\sigma^2 = \frac{n-1}{n}\frac{S^2}{\sigma^2}$ or better get (by Theorem 2.4.2 (iii)) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$. Form the confidence interval

$$\mathbb{P}\left\{\chi^2_{\alpha/2}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{1-\alpha/2}(n-1)\right\} = 1-\alpha$$

$$\mathbb{P}\left\{\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right\} = 1-\alpha$$

So given observed values $x_1, \ldots, x_n$ and let $s = \frac{1}{n-1}\sum_{i=1}^n (x_i - \overline{x})^2$,
then $\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}\right)$ is $100(1-\alpha)\%$ confidence interval for σ^2.

6.3 Two-Sample Problems

6.3.1 Two-sample independent Normal procedure

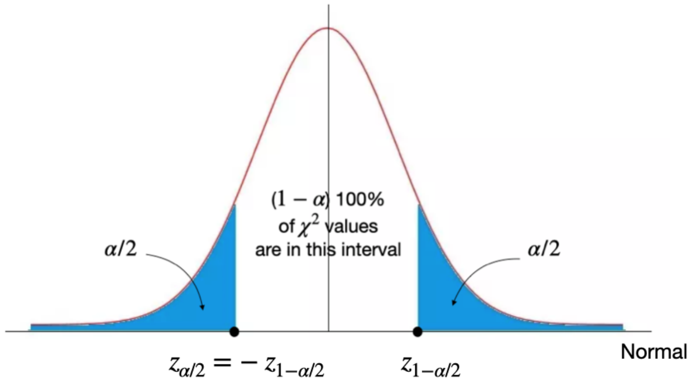
To know whether one population has a similar variance or mean than the other. Let $X_1, \ldots, X_{n_1}$ be a random sample from $\mathcal{N}(\mu_1, \sigma_1^2)$ and Let $Y_1, \ldots, Y_{n_2}$ be a random sample from $\mathcal{N}(\mu_2, \sigma_2^2)$. Assume X_i's and Y_i's are independent and denote $\overline{X}, \overline{Y}$ and S_1^2, S_2^2 are the sample means and sample variances of X_i's and Y_i's, respectively.

Procedure for variances

Recall $\frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \sim F(n_1-1, n_2-1)$ which provides a pivotal quantity for σ_2^2/σ_1^2. Form a confidence interval

$$\mathbb{P}\left\{f_{\alpha/2}(n_1-1, n_2-1) < \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} < f_{1-\alpha/2}(n_1-1, n_2-1)\right\} = 1-\alpha$$

$$\mathbb{P}\left\{\frac{S_2^2}{S_1^2} f_{\alpha/2}(n_1-1, n_2-1) < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2} f_{1-\alpha/2}(n_1-1, n_2-1)\right\} = 1-\alpha$$



and thus if s_1^2 and s_2^2 are estimates of S_1^2 and S_2^2, the $(1-\alpha)100\%$ confidence interval for σ_2^2/σ_1^2 is $\left(\frac{s_2^2}{s_1^2}f_{\alpha/2}(n_1-1,n_2-1), \frac{s_2^2}{s_1^2}f_{1-\alpha/2}(n_1-1,n_2-1)\right)$.

$\boxed{\textit{Example 6.3.1.}}$. Assume $n_1 = 16, n_2 = 21$ and $s_1^2 = 0.6$, $s_2^2 = 0.2$, and a 90% (or $\alpha = 0.05$) confidence interval is desired. One can find that $f_{0.05}(15,20) = 0.429$, $f_{0.95}(15,20) = 2.2$. It follows that $(0.2 \cdot 0.429/0.6 = 0.143, 0.2 \cdot 2.2/0.6 = 0.733)$ is a 90% confidence interval for σ_2^2/σ_1^2. Since the interval does not contain the value 1, we shall conclude that $\sigma_1^2 \neq \sigma_2^2$ or the two populations have different variance, and that only 10% of such conclusion will be incorrect.

Procedure for means

– **Variances are known** Assume σ_1^2, σ_2^2 are known. Recall $\overline{Y} - \overline{X} \sim \mathcal{N}(\mu_2 - \mu_1, \sigma_1^2/n_1 + \sigma_2^2/n_2)$ and it follows that $\frac{\overline{Y}-\overline{X}-(\mu_2-\mu_1)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim \mathcal{N}(0,1)$, which provides a pivotal quantity for $\mu_2 - \mu_1$. Form a confidence interval

$$\mathbb{P}\left\{-z_{1-\alpha/2} < \frac{\overline{Y} - \overline{X} - (\mu_2 - \mu_1)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < z_{1-\alpha/2}\right\} = 1 - \alpha$$

$$\mathbb{P}\left\{\overline{Y} - \overline{X} - z_{1-\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2} < \mu_2 - \mu_1 \right.$$
$$\left. < \overline{Y} - \overline{X} + z_{1-\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}\right\} = 1 - \alpha$$

and thus the $(1-\alpha)100\%$ confidence interval for $\mu_2 - \mu_1$ is

$$\left(\bar{y} - \bar{x} - z_{1-\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}, \bar{y} - \bar{x} + z_{1-\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}\right)$$

– **Variances are unknown but equal** In most cases the variances will be unknown, but in some cases it is reasonable to assume that the variances are unknown but equal. Assume $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and let $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$

$$\frac{(n_1+n_2-2)S_p^2}{\sigma^2} = \frac{(n_1-1)S_1^2}{\sigma^2} + \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_1+n_2-2)$$

Because $\frac{\overline{Y}-\overline{X}-(\mu_2-\mu_1)}{\sqrt{\sigma^2/n_1+\sigma^2/n_2}} \sim \mathcal{N}(0,1)$ and that we have showed that $\overline{X}, \overline{Y}$ are independent of S_1^2, S_2^2, by the definition of student's t distribtuion

$$\frac{\frac{\overline{Y}-\overline{X}-(\mu_2-\mu_1)}{\sqrt{\sigma^2/n_1+\sigma^2/n_2}}}{\sqrt{\frac{(n_1-1)S_1^2}{\sigma^2} + \frac{(n_2-1)S_2^2}{\sigma^2}}/\sqrt{n_1+n_2-2}} = \frac{\overline{Y}-\overline{X}-(\mu_2-\mu_1)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$$

which provides a pivotal quantity for $\mu_2 - \mu_1$. Form a confidence interval

$$\mathbb{P}\left\{-t_{1-\alpha/2}(n_1+n_2-2) < \frac{\overline{Y}-\overline{X}-(\mu_2-\mu_1)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} < t_{1-\alpha/2}(n_1+n_2-2)\right\}$$
$$= 1-\alpha$$
$$\mathbb{P}\left\{\overline{Y}-\overline{X} - t_{1-\alpha/2}(n_1+n_2-2)S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < \mu_2 - \mu_1\right.$$
$$\left. < \overline{Y}-\overline{X} + t_{1-\alpha/2}(n_1+n_2-2)S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right\} = 1-\alpha$$

and thus if s_1^2 and s_2^2 are estimates of S_1^2 and S_2^2, the $(1-\alpha)100\%$ confidence interval for $\mu_2 - \mu_1$ is

$$\left(\bar{y} - \bar{x} - t_{1-\alpha/2}(n_1+n_2-2)s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\right.$$

$$\overline{y} - \overline{x} + t_{1-\alpha/2}(n_1 + n_2 - 2)s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\Bigg)$$

$\boxed{\textit{Example 6.3.2.}}$. Assume $n_1 = 16, n_2 = 21$, $\overline{x} = 4.31, \overline{y} = 5.22$ and $s_1^2 = 0.12$, $s_2^2 = 0.1$. We might first check if the variances can be assumed to be equal. One can find that $f_{0.05}(15, 20) = 0.429$, $f_{0.95}(15, 20) = 2.2$. It follows that $(0.1 \cdot 0.429/0.12 = 0.358, 0.1 \cdot 2.2/0.12 = 1.83)$ is a 90% confidence interval for σ_2^2/σ_1^2, which contains the value 1. So we shall assume that $\sigma_1^2 = \sigma_2^2$.

Now the $s_p^2 = \frac{15 \cdot 0.12 + 20 \cdot 0.1}{35} = 0.109$ and suppose a 95% confidence interval is desired. We can find that $t_{0.975}(35) = 2.032$, so a 95% confidence interval for $\mu_2 - \mu_1$ is

$$\Big(5.22 - 4.31 - 2.032\sqrt{0.109}\sqrt{1/16 + 1/21} = 0.688,$$

$$5.22 - 4.31 - 2.032\sqrt{0.109}\sqrt{1/16 + 1/21} = 1.133\Big) = (0.688, 1.133)$$

– **Variances are unknown and unequal** For large-sample approximation, use

$$\frac{\overline{Y} - \overline{X} - (\mu_2 - \mu_1)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

For small-sample approximation, use

$$\frac{\overline{Y} - \overline{X} - (\mu_2 - \mu_1)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t\left(\frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}\right)$$

6.3.2 Paired sample Normal procedure

Let $X_1, \ldots, X_{n_1}$ be a random sample from $\mathcal{N}(\mu_1, \sigma_1^2)$ and Let $Y_1, \ldots, Y_{n_2}$ be a random sample from $\mathcal{N}(\mu_2, \sigma_2^2)$. Assume X_i's and Y_i's are dependent and the difference $D_i = Y_i - X_i \sim \mathcal{N}(\mu_2 - \mu_1, \sigma_D^2)$ with $\sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\sigma_{12}$. Let $\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i = \overline{Y} - \overline{X}$ and $S_D^2 = \sum_{i=1}^{n} \frac{(D_i - \overline{D})^2}{n-1}$. It follows from the definition of student's t distribution that

$$\sqrt{n}\frac{\overline{D} - (\mu_2 - \mu_1)}{S_D} \sim t(n-1)$$

which implies a $(1-\alpha)100\%$ confidence interval for $\mu_2 - \mu_1$ of

$$\left(\overline{d} - t_{1-\alpha/2}(n-1)s_D/\sqrt{n},\ \overline{d} + t_{1-\alpha/2}(n-1)s_D/\sqrt{n}\right)$$

6.3.3 Two-sample Binomial procedure

Suppose $X_1 \sim \mathcal{B}in(n_1, p_1)$ and $X_2 \sim \mathcal{B}in(n_2, p_2)$. Let $\hat{p}_1 = X_1/n_1$ and $\hat{p}_2 = X_2/n_2$, we can then use

$$\frac{\hat{p}_2 - \hat{p}_1 - (p_2 - p_1)}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \xrightarrow{d} \mathcal{N}(0,1)$$

to approximate the large sample confidence interval for $p_2 - p_1$.