# Lecture 16: Pivotal quantities

Another popular method of constructing confidence sets is the use of pivotal quantities defined as follows.

**Definition 9.2.6.**

A known function of $(X,\vartheta)$, $q(X,\vartheta)$, is called a pivotal quantity (or pivot) iff the distribution of $q(X,\vartheta)$ does not depend on any unknown quantity.

- A pivot is not a statistic, although its distribution is known.
- With a pivot $q(X,\vartheta)$, a level $1-\alpha$ confidence set for any given $\alpha$ can be obtained by finding a known Borel set $A$ (typically $A=[c_1,c_2]$) such that $P(q(X,\vartheta)\in A)\geq 1-\alpha$.
  Then a level $1-\alpha$ confidence set is $C(x)=\{\vartheta : q(x,\vartheta)\in A\}$, since
  $$\inf_{\theta\in\Theta} P\big(\vartheta\in C(X)\big) = \inf_{\theta\in\Theta} P\big(q(X,\vartheta)\in A\big) = P\big(q(X,\vartheta)\in A\big)\geq 1-\alpha$$
  If $q(X,\vartheta)$ has a continuous cdf, then we may choose $A$ such that $C(X)$ has confidence coefficient $1-\alpha$.

# Example (Fieller's interval).

Let $(X_{i1}, X_{i2})$, $i = 1, ..., n$, be iid bivariate normal with unknown $\mu_j = E(X_{1j})$, $\sigma_j^2 = \mathrm{Var}(X_{1j})$, $j = 1, 2$, and $\sigma_{12} = \mathrm{Cov}(X_{11}, X_{12})$.
Let $\vartheta = \mu_2/\mu_1$ be the parameter of interest ($\mu_1 \neq 0$).
$Y_1(\vartheta), ..., Y_n(\vartheta)$ are iid $N(0, \sigma_2^2 - 2\vartheta\sigma_{12} + \vartheta^2\sigma_1^2)$, $Y_i(\vartheta) = X_{i2} - \vartheta X_{i1}$.
Let

$$S^2(\vartheta) = \frac{1}{n-1}\sum_{i=1}^{n}[Y_i(\vartheta) - \bar{Y}(\vartheta)]^2 = S_2^2 - 2\theta S_{12} + \theta^2 S_1^2,$$

where $\bar{Y}(\vartheta)$ is the average of $Y_i(\vartheta)$'s and $S_i^2$ and $S_{12}$ are sample variances and covariance based on $X_{ij}$'s.
We know that $\sqrt{n}\bar{Y}(\vartheta)/S(\vartheta)$ has the t-distribution with $n-1$ degrees of freedom and, therefore, is a pivotal quantity.
Then

$$C(X) = \{\theta : n[\bar{Y}(\vartheta)]^2/S^2(\vartheta) \leq t_{n-1,\alpha/2}^2\}$$

is a confidence set for $\vartheta$ with confidence coefficient $1-\alpha$.
Note that $n[\bar{Y}(\vartheta)]^2 = t_{n-1,\alpha/2}^2 S^2(\vartheta)$ defines a parabola in $\vartheta$.
Depending on the roots of the parabola, $C(X)$ can be a finite interval, the complement of a finite interval, or the whole real line.

## Example.

We show an example of a confidence set for a two-dimensional parameter vector.

Let $X_1,...,X_n$ be iid from $N(\mu,\sigma^2)$ with unknown $\mu \in \mathscr{R}$ and $\sigma^2 > 0$.

Consider a confidence set for $\theta = (\mu,\sigma^2)$.

Since $(\bar{X},S^2)$ is sufficient and complete for $\theta$, we focus on $C(X)$ that is a function of $(\bar{X},S^2)$.

From Chapter 5, $\bar{X}$ and $S^2$ are independent and $(n-1)S^2/\sigma^2$ has the chi-square distribution with degrees of freedom $n-1$.

Since $\sqrt{n}(\bar{X}-\mu)/\sigma \sim N(0,1)$, a two dimensional pivot is

$$\left(\sqrt{n}(\bar{X}-\mu)/\sigma,\ (n-1)S^2/\sigma^2\right)$$

If $\tilde{c}_\alpha = \Phi^{-1}\left(\frac{1+\sqrt{1-\alpha}}{2}\right)$, then

$$P\left(-\tilde{c}_\alpha \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le \tilde{c}_\alpha\right) = \sqrt{1-\alpha}$$

Since the chi-square distribution is a known distribution, we can find two constants $c_{1\alpha}$ and $c_{2\alpha}$ such that

$$P\left(c_{1\alpha} \le \frac{(n-1)S^2}{\sigma^2} \le c_{2\alpha}\right) = \sqrt{1-\alpha}.$$

Then

$$P\left(-\tilde{c}_\alpha \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le \tilde{c}_\alpha, c_{1\alpha} \le \frac{(n-1)S^2}{\sigma^2} \le c_{2\alpha}\right) = 1-\alpha,$$

or

$$P\left(\frac{n(\bar{X}-\mu)^2}{\tilde{c}_\alpha^2} \le \sigma^2, \frac{(n-1)S^2}{c_{2\alpha}} \le \sigma^2 \le \frac{(n-1)S^2}{c_{1\alpha}}\right) = 1-\alpha.$$

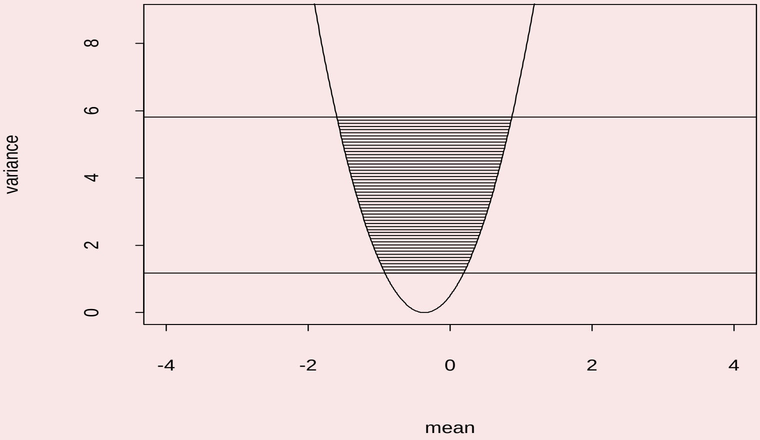
The left-hand side of the previous expression defines a set in the range of $\theta = (\mu, \sigma^2)$ bounded by two straight lines

$$\sigma^2 = (n-1)S^2/c_{i\alpha}, \qquad i = 1, 2,$$

and a curve

$$\sigma^2 = n(\bar{X}-\mu)^2/\tilde{c}_\alpha^2$$

See the shadowed part of the figure.




This set is a confidence set for $\theta$ with confidence coefficient $1-\alpha$, since the probability in the previous expression does not depend on any unknown parameter.

What happens if we replace $\sqrt{n}(\bar{X}-\mu)/\sigma$ by $\sqrt{n}(\bar{X}-\mu)/S$ in the pivot?

## Example 9.2.7 (location-scale pivots).

The approach of using pivots works well when the pdf of $X$ is in a location-scale family.

| Form of pdf | Type of pdf | Pivots |
|---|---|---|
| $f(x-\mu)$ | Location | $\bar{X}-\mu$ |
| $\sigma^{-1}f(x/\sigma)$ | Scale | $\bar{X}/\sigma$, $S^2/\sigma^2$, $X_{(n)}/\sigma$ |
| $\sigma^{-1}f((x-\mu)/\sigma)$ | Location-scale | $(\bar{X}-\mu)/S$, $S^2/\sigma^2$ |

The selection of a pivot depends on what parameter is of interest.
For example, $(\bar{X}-\mu)/\sigma$ is a pivot for getting a confidence set of the two-dimensional vector $\theta=(\mu,\sigma)$.
However, if we are interested in $\vartheta=\mu$, then $(\bar{X}-\mu)/\sigma$ is not the right pivot, although its distribution does not depend on anything unknown.
The right pivot is $(\bar{X}-\mu)/S$ or $\sqrt{n}(\bar{X}-\mu)/S$.
When $f$ is normal, the use of $\sqrt{n}(\bar{X}-\mu)/S$ gives the interval estimator

$$[\bar{X}-t_{n-1,\alpha/2}S/\sqrt{n},\ \bar{X}+t_{n-1,\alpha/2}S/\sqrt{n}]$$

If we also want a confidence interval for $\sigma^2$ or $\sigma$, then $(n-1)S^2/\sigma^2$ is a pivot, which leads to confidence intervals

$$\left[(n-1)S^2/b, (n-1)S^2/a\right] \qquad \text{and} \qquad \left[\sqrt{(n-1)S^2/b}, \sqrt{(n-1)S^2/a}\right]$$

where $a$ and $b$ are percentiles of the chi-square with degrees of freedom $n-1$.

When $f$ is $uniform(0,1)$, $\sigma^{-1}f(x/\sigma)$ is $uniform(0,\sigma)$, a scale family.
Note that $X_{(n)}/\sigma$ has pdf $nx^{n-1}$, $0 < x < 1$, and hence it is a pivot.
Using this pivot leads to an interval estimator $[aX_{(n)}, bX_{(n)}]$ for $\sigma$.
Since the coverage probability is $a^{-n} - b^{-n}$, setting $a^{-n} - b^{-n} = 1 - \alpha$ leads to a confidence coefficient $1 - \alpha$.

## Pivoting cdf's

Consider the situation where $\vartheta = \theta \in \Theta \subset \mathscr{R}$.
We can use a pivot based on the cdf of a real-valued sufficient statistic.

## Theorem 9.2.12 (pivoting a continuous cdf).

Let $T$ be a real-valued statistic with continuous cdf $F_\theta(t)$, where $\theta \in \Theta \subset \mathscr{R}$. Let $\alpha_1 + \alpha_2 = \alpha$, $0 < \alpha < 1$, be fixed nonnegative values.

If $F_\theta(t)$ is non-increasing in $\theta$ for each $t$, define $L(t)$ and $U(t)$ by

$$F_{U(t)}(t) = \alpha_1, \qquad F_{L(t)}(t) = 1 - \alpha_2$$

If $F_\theta(t)$ is non-decreasing in $\theta$ for each $t$, define $L(t)$ and $U(t)$ by

$$F_{U(t)}(t) = 1 - \alpha_2, \qquad F_{L(t)}(t) = \alpha_1$$

The interval estimator $[L(T),\ U(T)]$ has confidence coefficient $1 - \alpha$.

**Proof.**

The result follows from the fact that $F_\theta(T)$ has the uniform distribution of $(0,1)$ and thus a pivot, and if $\alpha_1 + \alpha_2 = \alpha$, then

$$P(\alpha_1 \leq F_\theta(T) \leq 1 - \alpha_2) = 1 - \alpha_2 - \alpha_1 = 1 - \alpha$$

Hence, we obtain a confidence set $\{\theta : \alpha_1 \leq F_\theta(T) \leq 1 - \alpha_2\}$.

- If some equations have no solution or multiple solutions, we define

  $$U(t) = \sup\{\theta : F_\theta(t) \geq \alpha_1\}, \qquad L(t) = \inf\{\theta : F_\theta(t) \leq 1 - \alpha_2\}$$

  when $F_\theta(t)$ is non-increasing, or, when $F_\theta(t)$ is non-decreasing,

  $$L(t) = \inf\{\theta : F_\theta(t) \geq \alpha_1\}, \qquad U(t) = \sup\{\theta : F_\theta(t) \leq 1 - \alpha_2\}$$

- For each equation, we only need to solve it for the value of $T = t$ that is actually observed. Even if one of the equation has no explicit solution, it can be solved numerically.
- If we only want a lower (or upper) confidence bound, then we can set one of the $\alpha_j$ to 0 and solve only one equation.
- If $F_\theta$ is from a family with MLR, then $F_\theta(t)$ is monotone in $\theta$ for every $t$.
- Even if $F_\theta(t)$ is not monotone in $\theta$, we can still apply this idea. But the resulting confidence set may not be an interval.

## Example 9.2.13.

Let $X_1, ..., X_n$ be iid with pdf $f_\theta(x) = e^{-(x-\theta)}$, $x > \theta$, where $\theta \in \mathscr{R}$.
The sufficient and complete statistic for $\theta$ is $T = X_{(1)}$ with pdf $ne^{-n(t-\theta)}$, $t > \theta$, which is from a family with MLR so that $F_\theta(t)$ is decreasing in $\theta$ for any $t$.
Then we define $L(t)$ and $U(t)$ by

$$\int_{U(t)}^{t} ne^{-n(u-U(t))} du = \alpha_1, \quad \int_{L(t)}^{t} ne^{-n(u-L(t))} du = 1 - \alpha_2$$

These integrals can be solved and

$$U(t) = t + n^{-1} \log(1-\alpha_1), \quad L(t) = t + n^{-1} \log \alpha_2$$

**Theorem 9.2.14 (pivoting a discrete cdf).**

Let $T$ be a real-valued discrete statistic with cdf $F_\theta(t)$, where $\theta \in \Theta \subset \mathscr{R}$. Let $\alpha_1 + \alpha_2 = \alpha$, $0 < \alpha < 1$, be fixed nonnegative values. If $F_\theta(t)$ is non-increasing in $\theta$ for each $t$, define $L(t)$ and $U(t)$ by

$$F_{U(t)}(t) = \alpha_1, \qquad P_{L(t)}(T \geq t) = \alpha_2$$

If $F_\theta(t)$ is non-decreasing in $\theta$ for each $t$, define $L(t)$ and $U(t)$ by

$$P_{U(t)}(T \geq t) = \alpha_2, \qquad F_{L(t)}(t) = \alpha_1$$

The interval estimator $[L(T), \, U(T)]$ has level $1 - \alpha$.

- The remarks after Theorem 9.2.12 still apply.
- The proof is very similar to the proof of Theorem 9.2.12, except that we do not obtain confidence coefficient $1-\alpha$ because $T$ is discrete.

## Example 9.2.15.

Let $X_1,...,X_n$ be iid random variables from *Poisson*$(\theta)$ with unknown $\theta > 0$ and $T = \sum_{i=1}^n X_i$, the sufficient and complete statistic for $\theta$.

Since $T \sim$ *Poisson*$(n\theta)$,

$$F_\theta(t) = \sum_{j=0}^{t} \frac{e^{-n\theta}(n\theta)^j}{j!}, \qquad t = 0,1,2,....$$

Since the Poisson family has monotone likelihood ratio in $T$ and $0 < F_\theta(t) < 1$ for any $t$, $F_\theta(t)$ is strictly decreasing in $\theta$. (The fact that $F_\theta(t)$ is strictly decreasing in $\theta$ can be directly proved.)

Also, $F_\theta(t)$ is continuous in $\theta$ and $F_\theta(t)$ tends to 1 and 0 as $\theta$ tends to 0 and $\infty$, respectively, and thus, Theorem 9.2.14 applies.

First, $U(t)$ is the unique solution of $F_\theta(t) = \alpha_1$ for any fixed $t$.

Since $P_\theta(T \geq t) = 1 - F_\theta(t-1)$ for $t > 0$, $L(t)$ is the unique solution of $F_\theta(t-1) = 1 - \alpha_2$ when $t > 0$.

When $t = 0$, $1 - F_\theta(-1) = 1$, we apply the remark after Theorem 9.2.12 and set $L(0) = 0$.

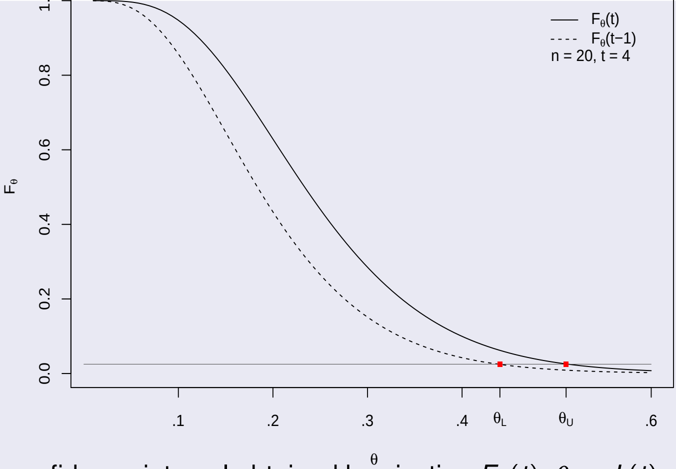



Figure: A confidence interval obtained by pivoting $F_\theta(t)$, $\theta_L = L(t)$, $\theta_U = U(t)$

In fact, in this case explicit forms of $L(t)$ and $U(t)$ can be obtained from the equality

$$\frac{1}{\Gamma(t)} \int_\lambda^\infty x^{t-1} e^{-x} dx = \sum_{j=0}^{t-1} \frac{e^{-\lambda} \lambda^j}{j!}, \qquad t = 1, 2, \ldots.$$

By $F_\theta(T) = \alpha_1$,

$$\alpha_1 = \sum_{k=0}^{t} \frac{e^{-n\theta}(n\theta)^k}{k!} = \frac{1}{\Gamma(t+1)} \int_{n\theta}^{\infty} x^t e^{-x} dx = P(\chi^2_{2(t+1)} \geq 2n\theta)$$

where $\chi^2_v$ is a random variable $\sim$ chi square with degrees of freedom $v$.
Similarly, $F_\theta(t-1) = 1 - \alpha_2$ gives

$$\alpha_2 = \sum_{k=t}^{\infty} \frac{e^{-n\theta}(n\theta)^k}{k!} = \frac{1}{\Gamma(t)} \int_0^{n\theta} x^{t-1} e^{-x} dx = P(\chi^2_{2t} \leq 2n\theta)$$

Hence,

$$U(t) = (2n)^{-1}\chi^2_{2(t+1),\alpha_1} \qquad L(t) = (2n)^{-1}\chi^2_{2t,1-\alpha_2},$$

where $\chi^2_{r,\alpha}$ is the $100(1-\alpha)$th percentile of the chi-square distribution with degrees of freedom $r$ and $\chi^2_{0,a}$ is defined to be 0.

## Gamma-Poisson relationship

To complete the argument, we need to show that if $X_\alpha \sim gamma(\alpha, \beta)$ with an integer $\alpha$ and $Y \sim Poisson(x/\beta)$, then

$$P(X_\alpha \leq x) = P(Y \geq \alpha), \quad x > 0$$

We use induction to prove this relationship.

- When $\alpha = 1$,

$$P(X_1 \le x) = \frac{1}{\beta}\int_0^x e^{-t/\beta}dt = e^{-t/\beta}\Big|_x^0 = 1 - e^{-x/\beta} = P(Y \ge 1)$$

- Assume that $P(X_{\alpha-1} \le x) = P(Y \ge \alpha - 1)$.
- From integration by parts,

$$\begin{aligned}
P(X_\alpha \le x) &= \frac{1}{(\alpha-1)!\beta^\alpha}\int_0^x t^{\alpha-1}e^{t/\beta}dt \\
&= \frac{1}{(\alpha-1)!\beta^{\alpha-1}}\left[t^{\alpha-1}e^{-t/\beta}\Big|_x^0 + (\alpha-1)\int_0^x t^{\alpha-2}e^{-t/\beta}dt\right] \\
&= \frac{1}{(\alpha-2)!\beta^{\alpha-1}}\int_0^x t^{\alpha-2}e^{-t/\beta}dt - \frac{x^{\alpha-1}e^{-x/\beta}}{(\alpha-1)!\beta^{\alpha-1}} \\
&= P(X_{\alpha-1} \le x) - \frac{e^{-x/\beta}}{(\alpha-1)!}\left(\frac{x}{\beta}\right)^{\alpha-1} \\
&= P(Y \ge \alpha - 1) - P(Y = \alpha - 1) = P(Y \ge \alpha)
\end{aligned}$$

## Example.

Let $X_1, ..., X_n$ be iid from a distribution with pdf $\theta x^{\theta-1}$, $0 < x < 1$, where $\theta > 0$ is unknown.

Note that $T = -\sum_{i=1}^n \log X_i$ is a complete and sufficient statistic for $\theta$, and $\theta T \sim$gamma(n,1).

Thus, $\theta T$ is a pivot and we can obtain a class of confidence intervals of the form $[c_1 T^{-1}, c_2 T^{-1}]$ with

$$\int_{c_1}^{c_2} f(x)dx = 1 - \alpha \qquad f \text{ is the pdf of } gamma(n,1)$$

to ensure the confidence coefficient $= 1 - \alpha$.

Consider testing hypotheses $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$.

By the result in Chapter 8, the acceptance region of a UMPU test is $A(\theta_0) = \{X : c_1 \leq \theta_0 T \leq c_2\}$, where $c_1$ and $c_2$ are determined by

$$\int_{c_1}^{c_2} f(x)dx = 1 - \alpha \qquad \text{and} \qquad \int_{c_1}^{c_2} xf(x)dx = n(1 - \alpha)$$

The 2nd equality ensures unbiasedness.

Thus, $[c_1 T^{-1}, c_2 T^{-1}]$ is a UMAU confidence interval.