

SECCIÓ DOCENT I PROBLEMES

La «**Secció docent i problemes**» té l'objectiu de publicar articles de caire docent, difícilment publicables en revistes de recerca. A cada número de *Qüestió* s'inclouran d'un a tres problemes i les solucions es donaran en el número següent.

Els lectors poden, si ho volen, proposar problemes amb les solucions pertinents i enviar-los a *Qüestió*, que farà una selecció i en publicarà els més adequats, fent la corresponent referència a l'autor.

També seran ben rebudes solucions alternatives a les propostes fetes per l'autor dels problemes; l'editorial es reservarà, però, el dret a publicar-les.

BOLSHEV'S METHOD OF CONFIDENCE LIMIT CONSTRUCTION

V. BAGDONAVIČIUS*

V. NIKOULINA**

M. NIKULIN*

Confidence intervals and regions for the parameters of a distribution are constructed, following the method due to L.N. Bolshev. This construction method is illustrated with Poisson, exponential, Bernouilli, geometric, normal and other distributions depending on parameters.

Keywords: Confidence limits, interval estimates.

* V. Bagdonavičius. University of Vilnius, Lithuania.

** V. Nikoulina. University of Bordeaux-2, France.

★ M. Nikulin. Steklov Mathematical Institute, St.Petersburg, Russia.

– Article rebut el maig de 1997.

– Acceptat l'octubre de 1997.

1. REGIONS, INTERVALS, CONFIDENCE LIMITS

Let $\mathbb{X} = (X_1, \dots, X_n)^T$ be a sample with realizations $x = (x_1, \dots, x_n)^T$, $x \in \mathcal{X} \subseteq R^n$. Suppose that X_i has a density $f(x; \theta)$, $\theta = (\theta_1, \dots, \theta_k)^T \in \Theta \subseteq R^k$, with respect to the Lebesgue measure,

$$H_0 : X_i \sim f(x; \theta), \quad \theta = (\theta_1, \dots, \theta_k)^T \in \Theta \subseteq R^k.$$

Let $b = b(\theta)$ be a function $b(\cdot) : \Theta \Rightarrow B \subseteq R^m$, B^0 is the interior of B

Definition 1 A random set $C(\mathbb{X})$, $C(\mathbb{X}) \subseteq B \subseteq R^m$ is called the confidence region for $b = b(\theta)$ with the confidence level γ ($0.5 < \gamma < 1$) if

$$\inf_{\theta \in \Theta} P_{\theta} \{C(\mathbb{X}) \ni b(\theta)\} = \gamma.$$

This definition implies for all $\theta \in \Theta$

$$P_{\theta} \{C(\mathbb{X}) \ni b(\theta)\} \geq \gamma.$$

In the case $b(\theta) \in B \subseteq R^1$ the confidence region is often an interval in R^1 ,

$$C(\mathbb{X}) =]b_i(\mathbb{X}), b_s(\mathbb{X})[\subseteq B \subseteq R^1,$$

and it is called the confidence interval with the confidence level γ for b . The statistics $b_i(\mathbb{X})$ and $b_s(\mathbb{X})$ are called the confidence limits of the confidence interval $C(\mathbb{X})$.

Definition 2 A statistic $b_i(\mathbb{X})$ ($b_s(\mathbb{X})$) is called the inferior (superior) confidence limit with the confidence level γ_1 (γ_2) (or inferior (superior) γ_1 (γ_2) - confidence limit briefly), if

$$\inf_{\theta \in \Theta} P_{\theta} \{b_i(\mathbb{X}) < b\} = \gamma_1 \left(\inf_{\theta \in \Theta} P_{\theta} \{b_s(\mathbb{X}) > b\} = \gamma_2 \right), \quad 0.5 < \gamma_j < 1$$

The $\gamma = 1 - \alpha$ confidence interval has the form $]b_i(\mathbb{X}), b_s(\mathbb{X})[$, where $b_i(\mathbb{X})$ and $b_s(\mathbb{X})$ are the $\gamma_1 = 1 - \alpha_1$ inferior and $\gamma_2 = 1 - \alpha_2$ superior confidence limits, respectively, such that $\alpha_1 + \alpha_2 = \alpha$, ($0 < \alpha_i < 0.5$). If $\alpha_1 = \alpha_2$, then take $\gamma_1 = \gamma_2 = 1 - \alpha/2$.

Definition 3 The intervals

$$\{b_i(\mathbb{X}), +\infty\} \quad \text{and} \quad \{-\infty, b_s(\mathbb{X})\}$$

are called the superior and inferior confidence intervals for b . Both intervals are unilateral.

2. THEOREM OF BOLSHEV

Lemma (Bolshev) Let $G(t)$ be the distribution function of the random variable T . Then for all $z \in [0, 1]$

$$(1) \quad P\{G(T) \leq z\} \leq z \leq P\{G(T-0) < z\}.$$

If T is continuous, then

$$P\{G(T) \leq z\} = z, \quad 0 \leq z \leq 1.$$

Proof: First, we prove the inequality

$$(2) \quad P\{G(T) \leq z\} \leq z, \quad 0 \leq z \leq 1.$$

If $z = 1$, then $P\{G(T) \leq 1\} \leq 1$. Fix $z \in [0, 1)$ and for this value of z consider the different cases.

1) There exists a solution y of the equation $G(y) = z$. Note

$$y_0 = \sup\{y : G(y) = z\}.$$

It can be:

a) $G(y_0) = z$. In this case

$$P\{G(T) \leq z\} \leq P\{T \leq y_0\} = G(y_0) = z.$$

b) $G(y_0) > z$. Then

$$P\{G(T) \leq z\} \leq P\{T < y_0\} = G(y_0 - 0) \leq z.$$

2) A solution of the equation $G(y) = z$ does not exist. In this case there exists y such that

$$G(y) > z \quad \text{et} \quad G(y-0) < z,$$

so

$$P\{G(T) \leq z\} \leq P\{T < y\} = G(y-0) < z.$$

The inequality (2) is proved.

We prove now the second inequality in (1) :

$$(3) \quad z \leq P\{G(T-0) < z\}, \quad 0 \leq z \leq 1.$$

Consider the statistic $-T$. Its distribution function is

$$G^-(y) = P\{-T \leq y\} = P\{T \geq -y\} = 1 - G(-y-0).$$

Replacing

$$T, z, G \text{ by } -T, 1-z \text{ and } G^-$$

in the inequality (2) we have:

$$P\{G^-(-T) \leq 1-z\} \leq 1-z, \quad 0 \leq z \leq 1.$$

This implies

$$P\{1 - G(T-0) \leq 1-z\} \leq 1-z,$$

$$P\{G(T-0) \geq z\} \leq 1-z,$$

$$P\{G(T-0) < z\} \geq z, \quad 0 \leq z \leq 1.$$

If T is continuous, then $G(t-0) = G(t)$, and (2) and (3) imply $P\{G(T) \leq z\} = z$ for all $z \in [0, 1]$.

The lemma is proved. ■

Theorem (Bolshev) *Suppose that the random variable $T = T(\mathbb{X}, b)$, $b \in B$, is such that its distribution function*

$$G(t; b) = P_{\theta}\{T \leq t\}$$

depends only on b for all $t \in R$ and the functions

$$I(b; x) = G(T(x, b) - 0; b) \quad \text{and} \quad S(b; x) = G(T(x, b); b)$$

are decreasing and continuous in b for all fixed $x \in X$. In this case:

1) *the statistic $b_i(\mathbb{X})$ such that*

$$(4) \quad b_i = b_i(\mathbb{X}) = \sup\{b : I(b; \mathbb{X}) \geq \gamma, b \in B\}, \quad \text{if this supremum exists,}$$

or

$$(5) \quad b_i = b_i(\mathbb{X}) = \inf B, \text{ otherwise}$$

is the inferior confidence limit for $b \in B^0$ with confidence level larger or equal to γ ;

2) the statistic $b_s(\mathbb{X})$ such that

$$(6) \quad b_s = b_s(\mathbb{X}) = \inf\{b : S(b; \mathbb{X}) \leq 1 - \gamma, \quad b \in B\}, \quad \text{if this infimum exists,}$$

or

$$(7) \quad b_s = b_s(\mathbb{X}) = \sup B, \text{ otherwise}$$

is the superior confidence limit for $b \in B^0$ with the confidence level larger or equal to γ ;

3) if $x \in \mathcal{X}$, is such that the functions $I(b; x)$ and $S(b; x)$ are strongly decreasing with respect to b , then $b_i(x)$ and $b_s(x)$ are the roots of the equations

$$(8) \quad I(b_i(x); x) = \gamma \quad \text{and} \quad S(b_s(x); x) = 1 - \gamma.$$

Proof: Denote $D = D(\mathbb{X})$ the event

$$D = \{\text{there exists } b \text{ such that } I(b; \mathbb{X}) \geq \gamma\}.$$

Then for the true value $b \in B^0$ we have (using Bolshev's lemma)

$$\begin{aligned} P\{b_i < b\} &= P\{(b_i < b) \cap D\} + P\{(b_i < b) \cap \bar{D}\} = \\ &= P\{((\sup b^* : I(b^*; \mathbb{X}) \geq \gamma, b^* \in B) < b) \cap D\} + P\{(\inf B < b) \cap \bar{D}\} = \\ &= P\{(I(b; \mathbb{X}) < \gamma) \cap D\} + P\{\bar{D}\} \geq P\{(I(b; \mathbb{X}) < \gamma) \cap D\} + P\{(I(b; \mathbb{X}) < \gamma) \cap \bar{D}\} = \\ &= P\{I(b; \mathbb{X}) < \gamma\} \geq \gamma. \end{aligned}$$

The theorem is proved. ■

Remark: Often, instead of the statistic T a sufficient statistic or some function of a sufficient statistic for a parameter b can be taken. □

3. EXAMPLES

1. Let $\mathbb{X} = (X_1, \dots, X_n)^T$ be a sample and suppose that X_i has a Poisson distribution with a parameter θ :

$$X_i \sim f(x; \theta) = \frac{\theta^x}{x!} e^{-\theta}, \quad x \in \mathcal{X} = \{0, 1, \dots\}, \quad \theta \in \Theta =]0, \infty[.$$

Denote

$$T = X_1 + \dots + X_n.$$

a) Show that the statistics

$$\theta_i = \frac{1}{2n} \chi_{1-\gamma_1}^2(2T) \quad \text{and} \quad \theta_s = \frac{1}{2n} \chi_{\gamma_2}^2(2T + 2)$$

are the inferior and superior confidence limits for θ with confidence levels larger or equal to γ_1 and γ_2 respectively; $\chi_{\alpha}^2(n)$ denotes the α -quantile of a chi-square distribution with n degrees of freedom.

b) Find a confidence interval for θ with confidence level larger or equal to γ .

Solution. The sufficient statistic T follows the Poisson distribution with parameter $n\theta$. Then

$$G(k; \theta) = P_{\theta}\{T \leq k\} = \sum_{i=0}^k \frac{(n\theta)^i}{i!} e^{-n\theta} = P\{\chi_{2k+1}^2 \geq 2n\theta\} = \mathcal{P}(2n\theta, 2k+2), \quad k = 0, 1, \dots$$

and

$$G(k-0; \theta) = P_{\theta}\{T < k\} = \sum_{i=0}^{k-1} \frac{(n\theta)^i}{i!} e^{-n\theta} = \mathcal{P}(2n\theta, 2k), \quad k = 1, 2, \dots,$$

$$G(k-0; \theta) = 0, \quad k = 0.$$

The functions I and S are

$$I(\theta; \mathbb{X}) = \begin{cases} \mathcal{P}(2n\theta, 2T), & \text{if } \mathbb{X} \neq 0, \\ 0, & \text{if } \mathbb{X} = 0, \end{cases}$$

$$S(\theta; \mathbb{X}) = \mathcal{P}(2n\theta, 2T + 2).$$

The function S is strictly decreasing for all T , $T \geq 0$, and I is strictly decreasing for all $T \neq 0$. In these cases the theorem of Bolshev implies (see (8)):

$$\mathcal{P}(2n\theta_i, 2T) = \gamma_1 \quad \mathcal{P}(2n\theta_s, 2T + 2) = 1 - \gamma_2,$$

from which it follows

$$(9) \quad \theta_i = \frac{1}{2n} \chi_{1-\gamma_1}^2(2T), \quad \theta_s = \frac{1}{2n} \chi_{\gamma_2}^2(2T + 2).$$

If $T = 0$ then $I(\theta; \mathbb{X}) = 0$. There is no such θ that

$$I(\theta; \mathbb{X}) = \gamma_1 > \frac{1}{2}.$$

The formula (5) implies

$$\theta_i = \inf_{\theta > 0} \theta = \inf]0, +\infty[= 0.$$

b) The interval $] \theta_i, \theta_s [$ is the confidence interval for θ with a confidence level larger or equal to $\gamma = 1 - \alpha$, if $\gamma_1 = 1 - \alpha_1$, $\gamma_2 = 1 - \alpha_2$, $\alpha_1 + \alpha_2 = \alpha$. If $\alpha_1 = \alpha_2$, take $\gamma_1 = \gamma_2 = 1 - \alpha/2$.

2. Let $\mathbb{X} = (X_1, \dots, X_n)^T$ be a sample and suppose that X_i has an exponential distribution with mean θ , $\theta > 0$:

$$(10) \quad X_i \sim f(x; \theta) = \frac{1}{\theta} \exp\left\{-\frac{x}{\theta}\right\} 1_{(x>0)}.$$

a) Find γ -confidence limits for θ .

b) Let $\mathbb{X}_n^{(r)} = (X_{(1)}, \dots, X_{(r)})^T$ be a type II censored sample from the distribution (10).

Find a γ -confidence interval for θ and the survival function

$$S(x; \theta) = P_\theta\{X_1 > x\}.$$

Solution. a). Denote

$$T = X_1 + \dots + X_n.$$

The sufficient statistic T follows a gamma distribution $G(n; \frac{1}{\theta})$ with parameters n and $1/\theta$:

$$P\{T \leq t\} = \frac{1}{(n-1)! \theta^n} \int_0^t u^{n-1} e^{-u/\theta} du, \quad t \geq 0,$$

and hence T/θ follows the gamma distribution $G(n; 1)$, and

$$\frac{2T}{\theta} = \chi_{2n}^2.$$

In this example the functions I and S can be taken as

$$I(\theta; \mathbb{X}) = S(\theta; \mathbb{X}) = 1 - \mathcal{P}\left(\frac{2T}{\theta}, 2n\right).$$

These functions are decreasing in θ and the formula (8) implies

$$1 - \mathcal{P}\left(\frac{2T}{\theta_i}, 2n\right) = \gamma \quad \text{and} \quad 1 - \mathcal{P}\left(\frac{2T}{\theta_s}, 2n\right) = 1 - \gamma,$$

from where we obtain

$$\frac{2T}{\theta_i} = \chi_{\gamma}^2(2n) \quad \text{and} \quad \frac{2T}{\theta_s} = \chi_{1-\gamma}^2(2n),$$

and hence

$$\theta_i = \frac{2T}{\chi_{\gamma}^2(2n)} \quad \text{and} \quad \theta_s = \frac{2T}{\chi_{1-\gamma}^2(2n)}.$$

b) As it is well known the statistic

$$T_r = \sum_{k=1}^r X_{(k)} + (n-r)X_{(r)}$$

follows a gamma distribution $G(r; \frac{1}{\theta})$, and hence the $\gamma = 1 - \alpha$ -confidence interval for θ is $] \theta_i, \theta_s [$, where

$$\theta_i = \frac{2T_r}{\chi_{1-\alpha/2}^2(2r)} \quad \text{and} \quad \theta_s = \frac{2T_r}{\chi_{\alpha/2}^2(2r)}.$$

Since the survival function $S(x; b) = e^{-x/\theta}$, $x > 0$, is increasing in θ , we have the γ -confidence interval $]S_i, S_s [$ for $S(x; \theta)$, where

$$S_i = e^{-x/\theta_i} \quad \text{and} \quad S_s = e^{-x/\theta_s}.$$

3. Let $\mathbb{X} = (X_1, \dots, X_n)^T$ be a sample from Bernoulli distribution with parameter θ :

$$X_i \sim f(x; \theta) = \theta^x (1 - \theta)^{1-x}, \quad x \in \mathcal{X} = \{0, 1\}, \quad \theta \in \Theta =]0, 1[.$$

Find the limits of confidence for θ with the confidence levels larger or equal to γ_1 .

Solution. It is clear that the sufficient statistic

$$T = \sum_{i=1}^n X_i$$

follows the binomial distribution $B(n, \theta)$ with parameters n and θ . Then

$$G(k; \theta) = P_{\theta}\{T \leq k\} = \sum_{i=0}^k \binom{n}{i} \theta^i (1 - \theta)^{n-i} =$$

$$I_{1-\theta}(n-k, k+1) = 1 - I_{\theta}(k+1, n-k), \quad k = 0, 1, \dots, n-1,$$

$$G(k; \theta) = 1, \quad \text{if } k = n,$$

where $I_x(a, b)$ is the beta distribution function with parameters a and b , and

$$G(k-0; \theta) = \sum_{i=0}^{k-1} \binom{n}{i} \theta^i (1 - \theta)^{n-i} =$$

$$1 - I_{\theta}(k, n-k+1), \quad k = 1, 2, \dots, n,$$

$$G(k-0; \theta) = 0, \quad \text{if } k = 0.$$

The functions I and S are

$$I(\theta; \mathbb{X}) = \begin{cases} I_{1-\theta}(n-T+1, T), & \text{if } T \neq 0 \\ 0, & \text{otherwise,} \end{cases}$$

$$S(\theta; \mathbb{X}) = \begin{cases} I_{1-\theta}(n-T, T+1), & \text{if } T \neq n \\ 1, & \text{if } T = n. \end{cases}$$

We remark that $S(\theta; \mathbb{X})$ is strictly decreasing in θ for $T \neq n$, and $I(\theta; \mathbb{X})$ is strictly decreasing in θ for $T \neq 0$, and hence from the formula (8) it follows that

$$I_{1-\theta_i}(n-T+1, T) = \gamma_1 \quad \text{for } T \neq 0$$

and

$$\theta_i = 0, \quad \text{if } T = 0,$$

$$I_{1-\theta_s}(n-T, T+1) = 1 - \gamma_1 \quad \text{for } T \neq n$$

and

$$\theta_s = 1, \quad \text{if } T = n.$$

Hence,

$$\theta_i = \begin{cases} 1 - x(\gamma_1; n-T+1, T), & \text{if } T \neq 0 \\ 0, & \text{if } T = 0, \end{cases}$$

$$\theta_s = \begin{cases} 1 - x(1 - \gamma_1; n-T, T+1), & \text{if } T \neq n \\ 1, & \text{if } T = n, \end{cases}$$

where $x(\gamma_1; a, b)$ is the γ_1 -quantile of the beta distribution with parameters a and b .

4. Let X be a discrete random variable with the cumulative distribution function

$$F(x; \theta) = P_\theta\{X \leq x\} = (1 - \theta^{\lfloor x \rfloor}) 1_{]0, +\infty[}(x), \quad x \in \mathbb{R}^1, \quad \theta \in \Theta =]0, 1[.$$

Find a γ -confidence interval for θ , if $X = 1$.

Solution. In this case

$$I(X; \theta) = F(X - 0; \theta) \quad \text{and} \quad S(X; \theta) = F(X; \theta).$$

If $X = 1$ then

$$I(1; \theta) = F(1 - 0; \theta) = F(0; \theta) = 0$$

and according to the formula (5) we have that the inferior confidence limit θ_i for θ with confidence level larger or equal to γ_1 is

$$\theta_i = \inf \theta = \inf]0, 1[= 0.$$

If $\gamma_1 = 1$ then $P\{\theta_i \leq \theta\} = \gamma_1$, so $\theta_i = 0$ is 1-confidence inferior limit for θ . On the other hand the function

$$S(1; \theta) = F(1; \theta) = 1 - \theta$$

is decreasing in θ , and hence according to the formula (8) we have

$$S(1; \theta_s) = 1 - \gamma_2,$$

from where $\theta_s = \gamma_2$, so the $\gamma_1 = 1$ and γ_2 confidence limits for θ are 0 and γ_2 , and a γ -confidence interval for θ is $]0, \gamma[$, since for $\gamma_1 = 1$ the equality $\gamma = \gamma_1 + \gamma_2 - 1$ is true when $\gamma_2 = \gamma$.

5. Let X_1 and X_2 be two independent random variables,

$$X_i \sim f(x; \theta) = e^{-(x-\theta)} 1_{[\theta, \infty[}(x), \quad \theta \in \Theta = \mathbb{R}^1.$$

Find the smallest γ -confidence interval for θ .

Solution. The likelihood function $L(\theta)$ for X_1 and X_2 is

$$L(\theta) = \exp\{-(X_1 + X_2 - 2\theta)\} 1_{[\theta, \infty[}(X_{(1)}),$$

from where it follows that $X_{(1)} = \min(X_1, X_2)$ is the minimal sufficient statistic for θ and $\hat{\theta} = X_{(1)}$ is the maximum of the function

$$l(\theta) = \ln L(\theta) = (2\theta - X_1 - X_2)1_{[\theta, \infty[}(X_{(1)}),$$

which is increasing in θ on the interval $]-\infty, X_{(1)}]$. Since for any $x \geq 0$

$$P_{\theta}\{X_{(1)} > x\} = P_{\theta}\{X_1 > x, X_2 > x\} = \left(\int_x^{\infty} e^{-(t-\theta)} dt \right)^2 = e^{-2(x-\theta)},$$

we have

$$P_{\theta}\{X_{(1)} \leq x\} = G(x; \theta) = \left(1 - e^{-2(x-\theta)}\right) 1_{[\theta, \infty[}(x), \quad x \in R^1.$$

In this example the functions $I(\theta; X_{(1)})$ and $S(\theta; X_{(1)})$ are

$$I(\theta; X_{(1)}) = S(\theta; X_{(1)}) = G(X_{(1)}; \theta) = 1 - e^{-2(X_{(1)}-\theta)}.$$

They are decreasing in θ and hence from the theorem of Bolshev we have

$$1 - e^{-2(X_{(1)}-\theta_i)} = \gamma_1, \quad \text{and} \quad 1 - e^{-2(X_{(1)}-\theta_s)} = 1 - \gamma_2,$$

thus

$$\theta_i = X_{(1)} + \frac{1}{2} \ln(1 - \gamma_1), \quad \text{and} \quad \theta_s = X_{(1)} + \frac{1}{2} \ln \gamma_2.$$

The interval $]\theta_i, \theta_s[$ is the γ -confidence interval for θ if $\gamma = \gamma_1 + \gamma_2 - 1$.

The length of this interval is

$$\theta_s - \theta_i = \frac{1}{2} [\ln \gamma_2 - \ln(1 - \gamma_1)].$$

We have to find γ_1 and γ_2 such that $\gamma_1 + \gamma_2 = 1 + \gamma$, $0.5 < \gamma_i \leq 1$ ($i = 1, 2$) and the interval $]\theta_i, \theta_s[$ is the shortest. We consider $\theta_s - \theta_i$ as the function of γ_2 . In this case

$$\begin{aligned} (\theta_s - \theta_i)' &= \frac{1}{2} [\ln \gamma_2 - \ln \gamma_2 - \gamma]' = \\ &= \frac{1}{2} \left(\frac{1}{\gamma_2} - \frac{1}{\gamma_2 - \gamma} \right) < 0, \end{aligned}$$

and hence $\theta_s - \theta_i$ is decreasing in γ_2 ($0.5 < \gamma_2 \leq 1$) and the minimal value of $\theta_s - \theta_i$ occurs when $\gamma_2 = 1$ and $\gamma_1 = 1 + \gamma - \gamma_2 = \gamma$. Since in this case

$$\theta_i = X_{(1)} + \frac{1}{2} \ln(1 - \gamma) \quad \text{and} \quad \theta_s = X_{(1)}$$

$$\min(\theta_s - \theta_i) = -\frac{1}{2} \ln(1 - \gamma) - \ln \sqrt{1 - \gamma}.$$

6. Let X_1 and X_2 be two independent random variables uniformly distributed on $[\theta - 1, \theta + 1]$, $\theta \in R^1$. Find the shortest γ -confidence interval for θ .

Solution. It is clear that $Y_i - \theta$ is uniformly distributed on $[-1, 1]$, from where it follows that the distribution of the random variable

$$T = X_1 + X_2 - 2\theta = Y_1 + Y_2$$

does not depend on θ . It is easy to show that

$$G(y) = P\{T \leq y\} = \begin{cases} 0, & y \leq -2, \\ \frac{1}{8}(y+2)^2, & -2 \leq y \leq 0, \\ 1 - \frac{(y-2)^2}{8}, & 0 \leq y \leq 2, \\ 1, & y \geq 2. \end{cases}$$

The function

$$G(T) = G(X_1 + X_2 - 2\theta), \theta \in R^1,$$

is decreasing in θ . From (8) it follows that the inferior and the superior confidence limits with the confidence levels γ_1 and γ_2 correspondingly ($0.5 < \gamma_i \leq 1$) satisfy the equations

$$G(X_1 + X_2 - 2\theta_i) = \gamma_1 \quad \text{and} \quad G(X_1 + X_2 - 2\theta_s) = 1 - \gamma_2,$$

from where we find

$$\theta_i = \frac{X_1 + X_2}{2} - 1 + \sqrt{2(1 - \gamma_1)} \quad \text{and} \quad \theta_s = \frac{X_1 + X_2}{2} + 1 - \sqrt{2(1 - \gamma_2)}.$$

It is easy to show that for given $\gamma = \gamma_1 + \gamma_2 - 1$ the function

$$\theta_s - \theta_i = 2 - \sqrt{2(1 - \gamma_1)} - \sqrt{2(1 - \gamma_2)}$$

has its minimal value (considered as function of γ_1 , $0.5 < \gamma_1 \leq 1$) when

$$\gamma_1 = \frac{1 + \gamma}{2}.$$

In this case $\gamma_2 = \frac{1 - \gamma}{2}$, so the shortest γ -confidence interval for θ is $]\theta_i, \theta_s[$ where

$$\theta_i = \frac{X_1 + X_2}{2} - 1 + \sqrt{1 - \gamma} \quad \text{and} \quad \theta_s = \frac{X_1 + X_2}{2} + 1 - \sqrt{1 - \gamma}.$$

7. Suppose that T is the number of shots until the first success. Find the $\gamma = 0.9$ confidence intervals for the probability p of success, if

a). $T = 1$; b). $T = 4$; c). $T = 10$.

Solution. The distribution of T is geometric :

$$P\{T = k\} = p(1-p)^{k-1}, \quad k = 1, 2, \dots$$

The values of the distribution function of T in the points k are

$$G(k; p) = \sum_{i=1}^k p(1-p)^{i-1} = 1 - (1-p)^k, \quad k = 1, 2, \dots$$

The functions I and S are

$$I(p; T) = 1 - (1-p)^{T-1}, \quad S(p; T) = 1 - (1-p)^T.$$

The functions $I(p; T)$ and $S(p; T)$ are increasing in p if $T > 1$ and $T \geq 1$, respectively. So they are decreasing in $q = 1 - p$.

It follows from the formula (8) that γ_1 lower and upper confidence limits satisfy the equations

$$\begin{aligned} 1 - q_i^{T-1} &= \gamma_1 \quad \text{for } T > 1, \\ 1 - q_s^T &= 1 - \gamma_1 \quad \text{for } T \geq 1. \end{aligned}$$

So

$$q_i = (1 - \gamma_1)^{\frac{1}{T-1}} \quad \text{for } T > 1, \quad q_s = \gamma_1^{\frac{1}{T}} \quad \text{for } T \geq 1$$

and

$$p_i = 1 - q_s = 1 - \gamma_1^{\frac{1}{T}} \quad \text{for } T \geq 1, \quad p_s = 1 - q_i = 1 - (1 - \gamma_1)^{\frac{1}{T-1}} \quad \text{for } T > 1.$$

If $T = 1$, then $q_i = \inf]0, 1[= 0$, $p_s = 1$.

To find the $\gamma = 1 - \alpha = 0.9$ confidence interval we take $\gamma_1 = 1 - \alpha/2 = \frac{1+\gamma}{2} = 0.95$.

So the $\gamma = 0.9$ confidence interval for p is (p_i, p_s) , where

$$\begin{aligned} p_i &= 0.05, \quad p_s = 1 \quad \text{for } T = 1, \\ p_i &= 1 - 0.95^{\frac{1}{4}} = 0.01274, \quad p_s = 1 - 0.05^{1/3} = 0.6316 \quad \text{for } T = 4, \\ p_i &= 1 - 0.95^{\frac{1}{10}} = 0.005116, \quad p_s = 1 - 0.05^{1/9} = 0.2831 \quad \text{for } T = 10. \end{aligned}$$

8. Let $\mathbb{X} = (X_1, \dots, X_n)^T$ be a sample and suppose that X_i has the normal distribution: $X_i \sim N(\mu, \sigma^2)$. Find a γ confidence interval for μ .

Solution. The sufficient statistic is (\bar{X}, S^2)

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Consider the statistic

$$T(\mathbb{X}, \mu) = \frac{\sqrt{n}(\bar{X} - \mu)}{S}.$$

The random variable $T(\mathbb{X}, \mu)$ has the Student distribution with $n - 1$ degrees of freedom and distribution function $F_{t_{n-1}}$. So

$$I(\mu, \mathbb{X}) = S(\mu, \mathbb{X}) = F_{t_{n-1}}(T(\mathbb{X}, \mu)).$$

The functions I and S are decreasing with respect to μ , so by the theorem of Bolshev

$$F_{t_{n-1}}(T(\mathbb{X}, \mu_i)) = \gamma_1 = \frac{1 + \gamma}{2}$$

$$F_{t_{n-1}}(T(\mathbb{X}, \mu_s)) = 1 - \gamma_1 = \frac{1 - \gamma}{2}$$

and

$$\mu_i = \bar{X} - \frac{S}{\sqrt{n}} t_{n-1} \left(\frac{1 + \gamma}{2} \right),$$

$$\mu_s = \bar{X} + \frac{S}{\sqrt{n}} t_{n-1} \left(\frac{1 + \gamma}{2} \right),$$

where $t_{n-1}(\alpha)$ is the α -quantile of the Student distribution with $n - 1$ degrees of freedom.

Confidence intervals for the variance, for the difference of two means, for the ratio of two variances, etc., can be obtained in a similar way.

REFERENCES

- [1] **Bolshev L.N.** (1965) «On the construction of confidence limits». *Theory of Prob. and its Applications*, **10**, 173-177.
- [2] **Bolshev L.N.** (1987) Selected papers. Theory of probability and mathematical statistics. Moscow, *Nauka*. p. 286.
- [3] **Bolshev L.N., Smirnov N.V.** (1983) Tables of mathematical statistics. Moscow, *Nauka*.

PROBLEMES PROPOSATS

PROBLEMA N° 66

Sean (X, Y) dos variables aleatorias discretas que toman los valores (x_i, y_j) , $i = 1, \dots, p$, $j = 1, \dots, q$, con probabilidad no nula. Sea P la matriz $p \times q$ con las densidades de probabilidad de Y condicionadas a X

$$P = \begin{pmatrix} f(y_1/x_1) \cdots f(y_q/x_1) \\ \vdots \\ f(y_1/x_p) \cdots f(y_q/x_p) \end{pmatrix}$$

Sea análogamente $Q = (f(x_i/y_j))$, $i = 1, \dots, p$, $j = 1, \dots, q$, la matriz $q \times p$ con las densidades de X condicionadas a Y .

Supongamos que $a_i(X)$, $i = 1, \dots, p$, $b_j(Y)$, $j = 1, \dots, q$, son asignaciones de los valores de las variables tales que

$$\mathbf{a} = (a_1(X), \dots, a_p(X))'$$

$$\mathbf{b} = (b_1(Y), \dots, b_q(Y))'$$

verifican

$$\mathbf{a} = \beta P \mathbf{b}$$

$$\mathbf{b} = \beta Q \mathbf{a}$$

siendo $\beta > 0$. Se pide:

1. Probar que \mathbf{a} es un vector propio de PQ de valor propio $\lambda = 1/\beta^2$.
2. Probar que $\lambda = 1$ es un valor propio de PQ de vector propio $\mathbf{1}_p = (1, \dots, 1)'$.
3. Probar que $0 \leq \lambda \leq 1$.

C.M. Cuadras

Universitat de Barcelona

PROBLEMA N° 67

Sea X una variable aleatoria con función de distribución absolutamente continua $F_\theta(X)$ y función de densidad $f_\theta(x)$, donde θ es un parámetro desconocido. Supongamos que $f_\theta(x)$ cumple las condiciones de regularidad (el soporte de X no depende del parámetro θ , es posible derivar bajo el signo integral, etc.)

Sea
$$I(\theta) = E \left\{ \left(\frac{\partial}{\partial \theta} \log f_\theta(X) \right)^2 \right\}$$

la cantidad de información de Fisher.

Probar que se verifica la desigualdad

$$4 \left(\frac{\partial}{\partial \theta} F_\theta(x) \right)^2 \leq I(\theta)$$

uniformemente en x .

C.M. Cuadras
Universitat de Barcelona