

**THE SHORTEST  
CONFIDENCE INTERVAL FOR THE WEIGHTED SUM  
OF TWO BINOMIAL PROPORTIONS**

**Stanisław Jaworski**

**Wojciech Zieliński**

Department of Econometrics and Statistics

Warsaw University of Life Sciences

Nowoursynowska 159, PL-02-787 Warszawa

e-mail: stanislaw\_jaworski@sggw.pl

e-mail: wojciech\_zielinski@sggw.pl

http: wojtek.zielinski.statystyka.info

**Summary.** Interval estimation of the probability of success in a Binomial model is considered. Zieliński (2018) showed that the confidence interval which uses information about non-homogeneity of the sample is better than the classical one. In the following paper the shortest confidence interval for non-homogenous sample is constructed.

**Keywords:** confidence interval, binomial distribution, binomial proportions

**AMS Classification:** 62F25

**Introduction.** Suppose that there are two suppliers of an item with unknown defectiveness  $\theta_1$  and  $\theta_2$ , respectively. It is known that the share of supply provided by the first supplier is  $w_1$ , while the share of the second one equals  $w_2 = 1 - w_1$ . It is of interest to estimate the overall defectiveness  $\vartheta = w_1\theta_1 + w_2\theta_2$  on the basis of a sample of size  $n$ .

The common approach (for example Decrouez & Robinson (2012)) is as follows: first we estimate  $\theta_1$  and  $\theta_2$  and then we construct a confidence interval for  $\vartheta = w_1\theta_1 + w_2\theta_2$ . Zieliński (2018) another approach applied. We want to construct a confidence interval for  $\vartheta$ , and we are not interested in estimating  $\theta_1$  and  $\theta_2$ . Note that for given  $\vartheta$  there are infinitely many  $\theta_1$  and  $\theta_2$  giving  $\vartheta$ . Hence averaging with respect to  $\theta_1$  and  $\theta_2$  is applied.

*ABin\_dwa\_przedzialy\_krotkie 13 stycznia 2019; 13:22; 1*

It was shown that the coverage probability of obtained confidence interval is at least the nominal confidence level. In what follows the shortest confidence interval is constructed. We confine ourselves to the case  $w_1, w_2 > 0$  and  $w_1 + w_2 = 1$ , because of its nice interpretation mentioned above.

**Confidence interval.** Suppose that  $n_1$  trials have been conducted with probability of success  $\theta_1$ , and  $n_2$  trials with probability of success  $\theta_2$ . We are interested in estimating of  $\vartheta = w_1\theta_1 + w_2\theta_2$  for known  $0 < w_1 < 1$  and  $w_2 = 1 - w_1$ . Assume that  $w_1 \leq w_2$ . Let  $\xi_1 \sim \text{Bin}(n_1, \theta_1)$ ,  $\xi_2 \sim \text{Bin}(n_2, \theta_2)$  and consider the random variable

$$\hat{\vartheta}_w = w_1 \frac{\xi_1}{n_1} + w_2 \frac{\xi_2}{n_2}.$$

Since we are interested in estimating  $\vartheta = w_1\theta_1 + w_2\theta_2$  on the basis of  $\hat{\vartheta}_w$ , we consider the new statistical model

$$(\mathcal{U}, \{\mathcal{P}(n_1, n_2, \vartheta), 0 \leq \vartheta \leq 1\}),$$

where

$$\mathcal{U} = \left\{ w_1 \frac{k_1}{n_1} + w_2 \frac{k_2}{n_2} : k_1 \in \{0, 1, \dots, n_1\}, k_2 \in \{0, 1, \dots, n_2\} \right\}.$$

The family  $\{\mathcal{P}(n_1, n_2, \vartheta), 0 \leq \vartheta \leq 1\}$  of distributions is as follows. Since for a given  $\vartheta \in (0, 1)$  the probability  $\theta_1$  is a number from the interval  $(a(\vartheta), b(\vartheta))$ , where

$$a(\vartheta) = \max \left\{ 0, \frac{\vartheta - w_2}{w_1} \right\}, \quad b(\vartheta) = \min \left\{ 1, \frac{\vartheta}{w_1} \right\} \quad \text{and} \quad L(\vartheta) = b(\vartheta) - a(\vartheta),$$

The probability of the event  $\{\hat{\vartheta}_w \leq u\}$  (for  $u \in \mathcal{U}$ ) equals (simply apply the law of total probability and averaging with respect to  $\theta_1$ )

$$P_\vartheta \{\hat{\vartheta}_w \leq u\} = \frac{1}{L(\vartheta)} \int_{a(\vartheta)}^{b(\vartheta)} \sum_{i_2=0}^{n_2} P_{\theta_1} \left\{ \xi_1 \leq \frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right) \right\} P_{\theta_2} \{\xi_2 = i_2\} d\theta_1. \quad (\star)$$

In the above integral  $\theta_2 = (\vartheta - w_1\theta_1)/w_2$ .

Note that the family  $\{\mathcal{P}(n_1, n_2, \vartheta), 0 \leq \vartheta \leq 1\}$  of distributions is decreasing in  $\vartheta$ , i.e. for a given  $u \in \mathcal{U}$ ,

$$P_{\vartheta_1} \{\hat{\vartheta}_w \leq u\} \geq P_{\vartheta_2} \{\hat{\vartheta}_w \leq u\} \quad \text{for} \quad \vartheta_1 < \vartheta_2.$$

It follows from the fact that the family of binomial distributions is decreasing in probability of a success and  $P_\vartheta \{\hat{\vartheta}_w \leq u\}$  is a convex combination of binomial distributions.

Let  $\hat{\vartheta}_w = u$  be observed. Let  $\gamma_1 \in [0, 1 - \gamma]$ . The confidence interval for  $\vartheta$  at confidence level  $\gamma$  is  $(\vartheta_L^w(u, \gamma_1), \vartheta_U^w(u, \gamma_1))$ , where

$$\begin{aligned} \vartheta_L^w(u, \gamma_1) &= \begin{cases} 0 & \text{for } u = 0, \\ \max\{\vartheta : P_\vartheta\{\hat{\vartheta}_w < u\} = \gamma + \gamma_1\} & \text{for } u > 0, \end{cases} \\ \vartheta_U^w(u, \gamma_1) &= \begin{cases} 1 & \text{for } u = 1, \\ \min\{\vartheta : P_\vartheta\{\hat{\vartheta}_w \leq u\} = \gamma_1\} & \text{for } u < 1. \end{cases} \end{aligned}$$

Note that

$$P_{\theta_1}\{\xi_1 < t\} = \begin{cases} P_{\theta_1}\{\xi_1 \leq t - 1\}, & \text{if } t \text{ is integer,} \\ P_{\theta_1}\{\xi_1 \leq t\}, & \text{elsewhere.} \end{cases}$$

For given  $\vartheta \in [0, 1]$ , the expected length of the confidence interval equals

$$l^w(\vartheta, \gamma_1) = \sum_{u \in \mathcal{U}} (\vartheta_U^w(u, \gamma_1) - \vartheta_L^w(u, \gamma_1)) g(u; \vartheta) \mathbf{1}_{(\vartheta_L^w(u, \gamma_1), \vartheta_U^w(u, \gamma_1))}(\theta),$$

where

$$\begin{aligned} g(u; \vartheta) &= P_\vartheta\{\hat{\vartheta}_w = u\} = P_\vartheta\left\{\frac{w_1}{n_1}\xi_1 + \frac{w_2}{n_2}\xi_2 = u\right\} \\ &= \frac{1}{L(\vartheta)} \int_{a(\vartheta)}^{b(\vartheta)} \sum_{i_2=0}^{n_2} P_{\theta_1}\left\{\xi_1 = \frac{n_1}{w_1}\left(u - w_2\frac{i_2}{n_2}\right)\right\} P_{\frac{\vartheta - w_1\theta_1}{w_2}}\{\xi_2 = i_2\} d\theta_1. \end{aligned}$$

The standard c.i. is obtained by taking  $\gamma_1 = (1 - \gamma)/2$ . Now the problem is in finding  $\gamma_1$  minimizing the length  $l^w(u, \gamma_1)$ .

Note that it may be assumed that  $u \leq 0.5$ . If  $u > 0.5$  the role of success and failure should be interchanged.

### Theorem 1.

1. For  $u \leq \max\left\{\frac{w_1}{n_1}, \frac{w_2}{n_2}\right\}$  the shortest confidence interval is one-sided.
2. In other cases the shortest confidence interval is two-sided.

**Lemma.** Let  $f, g : [0, 1] \rightarrow [0, 1]$  be two continuous functions such that  $f(0) = g(0) = 1$ ,  $f(1) = g(1) = 0$ ,  $f(x) > g(x)$  for all  $x \in (0, 1)$  and both functions are strictly decreasing. Let  $\gamma \in (0, 1)$  and let  $l(\gamma_1) = f^{-1}(\gamma_1) - g^{-1}(\gamma + \gamma_1)$  for  $\gamma_1 \in [0, 1 - \gamma]$ .

- a. If  $f'(0) = g'(0) = 0$  and  $f'(1) = g'(1) = 0$  then there exists  $\gamma_1^* \in (0, 1 - \gamma)$  such that  $l(\gamma_1^*) = \min\{l(\gamma_1) : \gamma_1 \in [0, 1 - \gamma]\}$ .
- b. If  $f'(0) = 0$ ,  $g'(0) < 0$  and  $f'(1) = g'(1) = 0$  then  $l(1 - \gamma) = \min\{l(\gamma_1) : \gamma_1 \in [0, 1 - \gamma]\}$ .

c. If  $f'(0) = g'(0) = 0$  and  $f'(1) < 0$ ,  $g'(1) = 0$  then  $l(0) = \min\{l(\gamma_1) : \gamma_1 \in [0, 1 - \gamma]\}$ .

*Proof of Lemma.* The derivative of  $l(\gamma_1)$  with respect to  $\gamma_1$  equals

$$\frac{dl(\gamma_1)}{d\gamma_1} = \frac{1}{f'(f^{-1}(\gamma_1))} - \frac{1}{g'(g^{-1}(\gamma + \gamma_1))}.$$

We have

$$\begin{cases} \gamma_1 \rightarrow 0 \Rightarrow f^{-1}(\gamma_1) \rightarrow f^{-1}(0) = 1 \text{ and } g^{-1}(\gamma + \gamma_1) \rightarrow g^{-1}(\gamma) \\ \gamma_1 \rightarrow 1 - \gamma \Rightarrow f^{-1}(\gamma_1) \rightarrow f^{-1}(1 - \gamma) \text{ and } g^{-1}(\gamma + \gamma_1) \rightarrow g^{-1}(1) = 0 \end{cases}$$

a. It is seen that

$$\gamma_1 \rightarrow 0 \Rightarrow \frac{dl(\gamma_1)}{d\gamma_1} \rightarrow -\infty \text{ and } \gamma_1 \rightarrow 1 - \gamma \Rightarrow \frac{dl(\gamma_1)}{d\gamma_1} \rightarrow +\infty.$$

Hence there exists  $\gamma_1^* \in (0, 1 - \gamma)$  such that  $\frac{dl(\gamma_1)}{d\gamma_1} \Big|_{\gamma_1^*} = 0$ . Since  $\frac{dl(\gamma_1)}{d\gamma_1} < 0$  for  $\gamma_1 < \gamma_1^*$  and  $\frac{dl(\gamma_1)}{d\gamma_1} > 0$  for  $\gamma_1 > \gamma_1^*$  hence we obtain thesis.

b. It is seen that

$$\gamma_1 \rightarrow 0 \Rightarrow \frac{dl(\gamma_1)}{d\gamma_1} < 0 \text{ and } \gamma_1 \rightarrow 1 - \gamma \Rightarrow \frac{dl(\gamma_1)}{d\gamma_1} \rightarrow -\infty.$$

Hence  $\frac{dl(\gamma_1)}{d\gamma_1} < 0$  for all  $\gamma_1 \in [0, 1 - \gamma]$ . Therefore we obtain thesis.

c. It is seen that

$$\gamma_1 \rightarrow 0 \Rightarrow \frac{dl(\gamma_1)}{d\gamma_1} \rightarrow +\infty \text{ and } \gamma_1 \rightarrow 1 - \gamma \Rightarrow \frac{dl(\gamma_1)}{d\gamma_1} > 0.$$

Hence  $\frac{dl(\gamma_1)}{d\gamma_1} > 0$  for all  $\gamma_1 \in [0, 1 - \gamma]$ . Therefore we obtain thesis.

*Proof of Theorem 1.*

For a given  $u \leq 0.5$  let  $G_u(\vartheta) = P_\vartheta\{\hat{\vartheta}_w < u\}$  and  $F_u(\vartheta) = P_\vartheta\{\hat{\vartheta}_w \leq u\}$ . We have

$$\vartheta_L^w(u, \gamma_1) = G_u^{-1}(\gamma + \gamma_1) \quad \text{and} \quad \vartheta_U^w(u, \gamma_1) = F_u^{-1}(\gamma_1).$$

The length of the confidence interval equals

$$l^w(u, \gamma_1) = \vartheta_U^w(u, \gamma_1) - \vartheta_L^w(u, \gamma_1) = F_u^{-1}(\gamma_1) - G_u^{-1}(\gamma + \gamma_1).$$

Let

$$U(\vartheta; u, \theta_1) = \sum_{i_2=0}^{n_2} P_{\theta_1} \left\{ \xi_1 \leq \frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right) \right\} P_{\frac{\vartheta-w_1\theta_1}{w_2}} \{ \xi_2 = i_2 \}$$

and

$$L(\vartheta; u, \theta_1) = \sum_{i_2=0}^{n_2} P_{\theta_1} \left\{ \xi_1 < \frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right) \right\} P_{\frac{\vartheta-w_1\theta_1}{w_2}} \{ \xi_2 = i_2 \}.$$

We have

$$F_u(\vartheta) = \begin{cases} \frac{w_1}{\vartheta} \int_0^{\vartheta/w_1} U(\vartheta; u, \theta_1) d\theta_1, & \vartheta < w_1, \\ \int_0^1 U(\vartheta; u, \theta_1) d\theta_1, & w_1 < \vartheta < w_2, \\ \frac{w_1}{1-\vartheta} \int_{\frac{\vartheta-w_2}{w_1}}^1 U(\vartheta; u, \theta_1) d\theta_1, & \vartheta > w_2 \end{cases}$$

and

$$G_u(\vartheta) = \begin{cases} \frac{w_1}{\vartheta} \int_0^{\vartheta/w_1} L(\vartheta; u, \theta_1) d\theta_1, & \vartheta < w_1, \\ \int_0^1 L(\vartheta; u, \theta_1) d\theta_1, & w_1 < \vartheta < w_2, \\ \frac{w_1}{1-\vartheta} \int_{\frac{\vartheta-w_2}{w_1}}^1 L(\vartheta; u, \theta_1) d\theta_1, & \vartheta > w_2. \end{cases}$$

The derivative  $dF_u(\vartheta)/d\vartheta$  for  $0 < \vartheta < w_1$  equals

$$\begin{aligned} & \int_0^1 \frac{dU(\vartheta; u, \frac{\vartheta}{w_1} \theta_1)}{d\vartheta} d\theta_1 = \\ & - \int_0^1 \sum_{i_2=0}^{n_2} \binom{n_2}{i_2} \left( \frac{\vartheta(1-\theta_1)}{w_2} \right)^{i_2} \left( \frac{w_2 - \vartheta(1-\theta_1)}{w_2} \right)^{n_2-i_2} \\ & \left( \frac{\left( \frac{\vartheta\theta_1}{w_1} \right)^{\frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right)} \left( 1 - \frac{\vartheta\theta_1}{w_1} \right)^{n_1 - \frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right) - 1}}{\text{Beta} \left( n_1 - \frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right), \frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right) + 1 \right)} + \right. \\ & \left. \frac{w_1(i_2 w_2 - n_2 \vartheta(1-\theta_1)) B \left( n_1 - \frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right), \frac{n_1}{w_1} \left( u - \frac{w_2}{n_2} i_2 \right) + 1; 1 - \frac{\vartheta\theta_1}{w_1} \right)}{\vartheta(1-\theta_1)(w_2 - \vartheta(1-\theta_1))} \right) d\theta_1. \end{aligned}$$

Here  $B(\cdot, \cdot; \cdot)$  is the cumulative distribution function of the beta distribution.

It is seen that

$$\lim_{\vartheta \rightarrow 0} dF_u(\vartheta)/d\vartheta = 0.$$

1. Let  $\frac{w_1}{n_1} = \max \left\{ \frac{w_1}{n_1}, \frac{w_2}{n_2} \right\}$ . Then

$$L(\vartheta; u, \theta_1) = B \left( n_1 - 1, 1; 1 - \theta_1 \right) \left( 1 - \frac{\vartheta - w_1\theta_1}{w_2} \right)^{n_2}$$

and  $dG_u(\vartheta)/d\vartheta$  for  $0 < \vartheta < w_1$  equals

$$-w_1 \frac{\Gamma(n_2 + 1)}{\vartheta(1 - \vartheta)} \left( \left(1 - \frac{\vartheta}{w_1}\right)^{n_1+1} {}_2F_1 \left[ 1, n_1 + n_2 + 1; n_2 + 1; \frac{w_2}{1 - \vartheta} \right] - \left(1 - \frac{\vartheta}{w_2}\right)^{n_2} {}_2F_1 \left[ 1, n_1 + n_2 + 1; n_2 + 1; \frac{w_2 - \vartheta}{1 - \vartheta} \right] \right),$$

where

$${}_2F_1 [x, y; z; t] = \sum_{j=0}^{\infty} \frac{(x)_j (y)_j}{(z)_j} \frac{t^j}{j!}$$

is the confluent hypergeometric function.

It is seen that

$$\lim_{\vartheta \rightarrow 0} dG_u(\vartheta)/d\vartheta < 0.$$

Applying point **b.** of Lemma we obtain that the shortest c.i. for  $\vartheta$  is one-sided.

For  $\frac{w_2}{n_2} = \max \left\{ \frac{w_1}{n_1}, \frac{w_2}{n_2} \right\}$  the proof is similar. In formula  $(\star)$  the role of  $\theta_1$  and  $\theta_2$  should be interchanged.

**2.** It is easy to check that in other cases  $\lim_{\vartheta \rightarrow 0} dG_u(\vartheta)/d\vartheta = 0$ . From point **a.** of Lemma it follows that the shortest c.i. is two-sided.

For  $u > 0.5$  we have

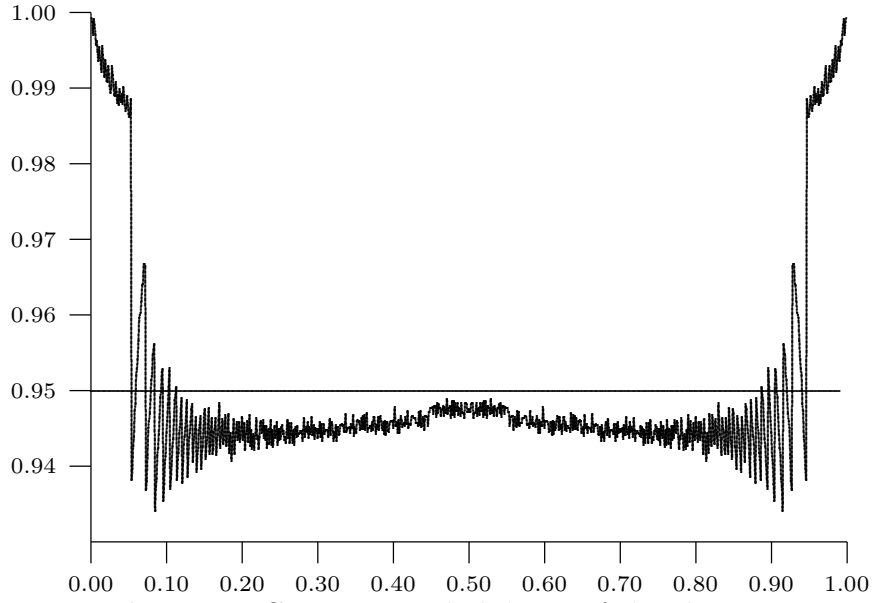
$$\vartheta_U^w(u, \gamma_1) = 1 - \vartheta_L^w(1 - u, (1 - \gamma) - \gamma_1) \quad \text{and} \quad \vartheta_L^w(u, \gamma_1) = 1 - \vartheta_U^w(1 - u, (1 - \gamma) - \gamma_1). \quad (\star\star)$$

**Theorem 2.** For  $u = 0.5$  the standard c.i. is the shortest one.

*Proof.*

From  $(\star\star)$  it is seen  $\vartheta_L^w(0.5, \gamma_1) = \vartheta_L^w(0.5, (1 - \gamma) - \gamma_1)$ . The equality is fulfilled for  $\gamma_1 = (1 - \gamma)/2$ .

For  $n_1 = 20$ ,  $n_2 = 30$ ,  $w_1 = 0.30$  and  $\gamma = 0.95$  the coverage probability of the shortest c.i. is shown in Figure 1. Note that for some probabilities  $\vartheta$  the coverage probability is smaller than the nominal confidence level. This is in contradiction with the definition of the confidence interval (see Neyman (1934), Cramér (1946), Lehmann (1959), Silvey (1970)).



**Figure 1.** Coverage probabilities of the shortest c.i.  
for  $n_1 = 20$ ,  $n_2 = 30$ ,  $w_1 = 0.30$

### Randomized c.i.

To avoid the above mentioned disadvantage a randomization is introduced (c.f. Zieliński 2017). Let

$$\mathcal{U} = \left\{ w_1 \frac{k_1}{n_1} + w_2 \frac{k_2}{n_2} : k_1 = 0, 1, \dots, n_1, k_2 = 0, 1, \dots, n_2 \right\}.$$

For  $u \in \mathcal{U}$  let

$$u^- = \max \{v < u : v \in \mathcal{U}\}, \quad u^+ = \min \{w > u : w \in \mathcal{U}\}.$$

Suppose that  $\hat{v}_w = u$  is observed. Consider two r.v.'s

$$\eta_d \sim U(0, u - u^-) \text{ and } \eta_s \sim U(0, u^+ - u)$$

and let  $\tilde{v}_s = \hat{v}_w + \eta_s$  and  $\tilde{v}_d = \hat{v}_w - \eta_d$ . Distributions of  $\tilde{v}_s$  and  $\tilde{v}_d$  are easy to obtain:

$$P_\theta\{\tilde{v}_s \leq t\} = P_\theta\{\hat{v}_w \leq [t]^- \} + \frac{[t] - [t]^-}{[t] - [t]^-} P_\theta\{\hat{v}_w = [t]\},$$

$$P_\theta\{\tilde{v}_d \leq t\} = P_\theta\{\hat{v}_w \leq [t]\} + \frac{[t]}{[t]^+ - [t]} P_\theta\{\hat{v}_w = [t]^+\},$$

where

$$[t] = \max \{v \leq t : v \in \mathcal{U}\}, \quad [t]^- = \max \{v < [t] : v \in \mathcal{U}\}, \quad [t]^+ = \min \{v > [t] : v \in \mathcal{U}\},$$

$$[t] = t - [t], [t]^- = -1 \text{ for } [t] = 0, [t]^+ = 2 \text{ for } [t] = 1.$$

The shortest confidence interval  $(\vartheta_L, \vartheta_U)$  at the confidence level  $\gamma$  for observed  $\hat{\vartheta}_w = u$ ,  $\eta_d = y$  and  $\eta_s = t$  will be obtained as a solution with respect to  $\vartheta$  of the following problem:

$$\begin{cases} \vartheta_U - \vartheta_L = \min! \\ P_{\vartheta_L} \left\{ \hat{\vartheta}_w \leq u^- \right\} + \frac{t}{u-u^-} P_{\vartheta_L} \left\{ \hat{\vartheta}_w = u \right\} = \gamma_2, \\ P_{\vartheta_U} \left\{ \hat{\vartheta}_w \leq u \right\} + \frac{y}{u^+-u} P_{\vartheta_U} \left\{ \hat{\vartheta}_w = u^+ \right\} = \gamma_1, \\ \gamma_2 - \gamma_1 = \gamma. \end{cases}$$

Note that the randomization may be simplified in the following way. Let  $\eta$  be a r.v. distributed as  $U(0, 1)$ . Than for observed  $\hat{\vartheta}_w = u$  and  $\eta = y$  the ends of the shortest randomized c.i. are solutions of

$$\begin{cases} \vartheta_U - \vartheta_L = \min! \\ P_{\vartheta_L} \left\{ \hat{\vartheta}_w \leq u^- \right\} + y P_{\vartheta_L} \left\{ \hat{\vartheta}_w = u \right\} = \gamma_2, \\ P_{\vartheta_U} \left\{ \hat{\vartheta}_w \leq u \right\} + y P_{\vartheta_U} \left\{ \hat{\vartheta}_w = u^+ \right\} = \gamma_1, \\ \gamma_2 - \gamma_1 = \gamma \end{cases}$$

or equivalently

$$\begin{cases} \vartheta_U - \vartheta_L = \min! \\ P_{\vartheta_L} \left\{ \hat{\vartheta}_w < u \right\} + y P_{\vartheta_L} \left\{ \hat{\vartheta}_w = u \right\} = \gamma_2, \\ P_{\vartheta_U} \left\{ \hat{\vartheta}_w < u^+ \right\} + y P_{\vartheta_U} \left\{ \hat{\vartheta}_w = u^+ \right\} = \gamma_1, \\ \gamma_2 - \gamma_1 = \gamma. \end{cases}$$

For  $u = 0$  we take  $\vartheta_L = 0$  and for  $u = 1$  we take  $\vartheta_U = 1$ .

**Theorem 3.** For  $u > 0$  there exists the shortest randomized c.i.

*Proof.*

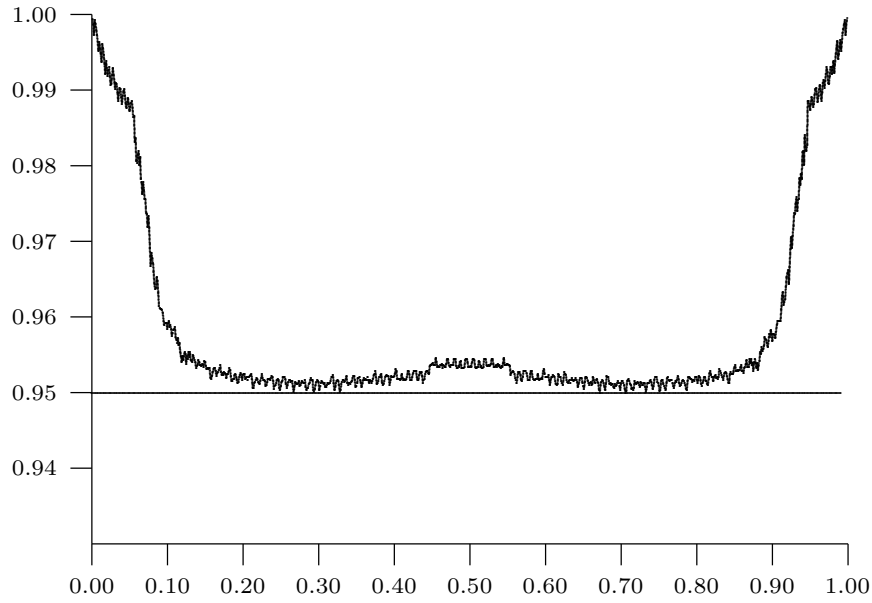
The proof is similar to the proof of Theorem 1.

For  $u > 0.5$  we have

$$\vartheta_U^w(u, \gamma_1) = 1 - \vartheta_L^w(1-u, 1-y, (1-\gamma) - \gamma_1) \quad \text{and} \quad \vartheta_L^w(u, y, \gamma_1) = 1 - \vartheta_U^w(1-u, 1-y, (1-\gamma) - \gamma_1)$$

In Figure 2 the coverage probability of the shortest c.i. for  $n_1 = 20$ ,  $n_2 = 30$ ,  $w_1 = 0.30$  and  $\gamma = 0.95$  is shown. It is seen that for all  $\vartheta$  the coverage probability is at least the nominal confidence level.





**Figure 2.** Coverage probabilities of the randomized shortest c.i. for  $n_1 = 20$ ,  $n_2 = 30$ ,  $w_1 = 0.30$

### An example

Let  $w_1 = 0.3$ . Consider an experiment consisting of  $n_1 = 20$  and  $n_2 = 30$  Bernoulli trials in which  $\hat{\vartheta}_w = 0.03$  were observed. Let  $\gamma = 0.95$ . The standard confidence interval takes on the form

$$(0.004672159, 0.111730732).$$

The length of that confidence interval equals 0.107058574.

To calculate the randomized shortest confidence interval one has to draw a value  $y$  of the auxiliary variable  $\eta$  and then calculate the ends of the confidence interval. The uniform random number generator gives  $y = 0.2$  and the randomized shortest confidence interval takes on the form

$$(0.000883203, 0.097443898).$$

The length of that confidence interval is 0.096560695. The length of the proposed confidence interval equals 90% of the length of the standard confidence interval.

The final report may look as follows:

$$w_1 = 0.3, n_1 = 20, n_2 = 30, \hat{\vartheta}_w = 0.03, y = 0.2, \gamma = 0.95,$$

$$\vartheta \in (0.000883203, 0.097443898).$$

## Conclusions

In practical applications it is important to have conclusions as precise as possible. Hence the use of the randomized shortest confidence intervals is recommended, especially for small sample sizes. Those intervals are very easy to obtain with the aid of the standard computer software (see Appendix).

## References

- CLOPPER, C. J. & PEARSON, E. S. (1934), The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika* 26, 404-413
- CRAMÉR, H. (1946). *Mathematical Methods in Statistics*, Princeton University Press (Nineteenth printing 1999)
- DECROUEZ, G. & ROBINSON, A. P. (2012) Confidence intervals for the weighted sum of two independent binomial proportions, *Aust. N. Z. J. Stat.* 54(3), 281-299
- LEHMANN, E. L. (1959). *Testing Statistical Hypothesis*, Springer (Third edition 2005)
- NEYMAN, J. (1934). On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. *Journal of the Royal Statistical Society* 97, 558-625
- SILVEY, S. D. (1970). *Statistical Inference*, Chapman & Hall (Eleventh edition 2003)
- ZIELIŃSKI W. (2011). Comparison of confidence intervals for fraction in finite populations, *Quantitative Methods in Economy XII*, 177-182.
- ZIELIŃSKI, W. (2016). A remark on estimating defectiveness in sampling acceptance inspection. *Colloquium Biometricum* 46, 9-14.
- ZIELIŃSKI, W. (2017) The shortest Clopper-Pearson randomized confidence interval for binomial probability. *REVSTAT-Statistical Journal* 15(1), 141-153.
- ZIELIŃSKI, W. (2018) Confidence interval for the weighted sum of two Binomial proportions. *Applicationes Mathematicae* 45(1), 53-60, doi: 10.4064/am2349-12-2017
- ZIELIŃSKI, W. (2019) A new exact confidence interval for the difference of two Binomial proportions. *REVSTAT-Statistical Journal*, (to appear).

## Appendix

An exemplary R code for calculating the shortest confidence interval is enclosed.

```
library(lpSolve)
FFbin=function(m,n,q){pbinom(m,n,q)} #Binomial CDF
fbin=function(m,n,q){aaa=(m==floor(m)); aaa*dbinom(floor(m),n,q)+(1-aaa)*0} #Binomial PDF
Fbin=function(m,n,q){war=(m==floor(m)); war*FFbin(m-1,n,q)+(1-war)*FFbin(m,n,q)}
#inputs
confllevel=0.95
w1=0.3 #share of strata1
n1=20 #sample1 size
n2=30 #sample2 size
k1=2 #no of succeses in sample1
k2=10 #no of succeses in sample2
#end of inputs
w2=1-w1
wz=w1*k1/n1+w2*k2/n2
los=runif(1, min = 0, max = 1)
iksy=c(0:n2)
smallnumber=1e-13
#important functions
#pdf at xx
probab=function(kk,mm,qqq){
integrate(Vectorize(function(ppp){fbin(kk,n1,ppp)*fbin(mm,n2,(qqq-w1*ppp)/w2})),
max(0,(qqq-w2)/w1),min(1,qqq/w1))$value/(min(1,qqq/w1)-max(0,(qqq-w2)/w1))}
density=function(xx,qqq){
gest=0
for(ii in iksy){gest=gest+probab(round((n1/w1)*(xx-(w2/n2)*ii),2),ii,qqq)}
gest}
#cdf at xx
distrib=function(xx,qqq){
integrate(Vectorize(function(ppp){
sum(FFbin(round((n1/w1)*(xx-(w2/n2)*iksy),2),n1,ppp)*fbin(iksy,n2,(qqq-w1*ppp)/w2)})),
max(0,(qqq-w2)/w1),min(1,qqq/w1))$value/(min(1,qqq/w1)-max(0,(qqq-w2)/w1))}
#randomized cdf
randomized=function(xx,vv,yy,qqq){distrib(xx,qqq)+yy*density(vv,qqq)}
#end of c.i.
TheEnd = function(xx,vv,yy,prawd)
{uniroot(function(t) {randomized(xx,vv,yy,t) - prawd}, lower = smallnumber, upper =1-smallnumber, tol =
1e-20)$root}
Leng = function(xx,ww,vv,yy,q,s){TheEnd(ww,vv,yy,s)-TheEnd(xx,ww,yy,q+s)}
FindMinimumLeng = function(ww,xx,vv,yy,q){
a = 0; b = 1-confllevel; #minimum in (a,b)
wspolczynnik = (sqrt(5) - 1)/2; #golden ratio
xL = b - wspolczynnik*(b - a); #left sampling point
xR = a + wspolczynnik*(b - a); #right sampling point
epsi = 10-10;
while((b- a) > epsi)
{
if(Leng(ww,xx,vv,los,confllevel,xL)<Leng(ww,xx,vv,los,confllevel,xR))
{b = xR; xR = xL; xL = b - wspolczynnik*(b - a);}
else
```

```

{a = xL; xL = xR; xR = a + wspolczynnik*(b - a);}
}
amin = (a + b)/2;
amin
}
#end of definitions
#determining uuplus and uuminus
eps=0.0001
obj.fun = c(w1/n1, w2/n2)
constr = matrix(c(w1/n1, w2/n2, 1, 0, 0, 1) , nrow = 3, byrow = TRUE)
constr.dir = c("<","<=","<=")
rhs = c(wz-eps,n1,n2)
prod.sol = lp("max", obj.fun, constr, constr.dir, rhs, int.vec=1:2)
uum=prod.sol$objval #objective function value
constr.dir = c(">","<=","<=")
rhs = c(wz+eps,n1,n2)
prod.sol = lp("min", obj.fun, constr, constr.dir, rhs, int.vec=1:2)
uup=prod.sol$objval #objective function value
uu=round(wz,8)
uum=round(uum,8)
uup=round(uup,8)
smin=FindMinimumLeng(uum,uu,uup,los,conflvel)
LowerEnd=if(uum<=0) 0 else TheEnd(uum,uu,los,conflvel+smin)
UpperEnd=if(uup>=1) 1 else TheEnd(uu,uup,los,smin)
#output
cat("weight of strata 1:", w1,"\n",
"sample 1 size:", n1,"\n",
"sample 2 size:", n2,"\n",
"no of successes in sample 1:", k1,"\n",
"no of successes in sample 2:", k2,"\n",
"estimate of theta:", uu,"\n",
"random number:", format(los,digits=3),"\n",
"left tail:", format(smin,digits=20),"\n",
"left end of the shortest confidence interval:", format(LowerEnd,digits=20,scientific=FALSE),"\n",
"right end of the shortest confidence interval:", format(UpperEnd,digits=20,scientific=FALSE),"\n")

```