UNIFORM ASYMPTOTIC NORMALITY FOR THE BERNOULLI SCHEME

Wojciech Niemiro

Faculty of Mathematics and Computer Science
Nicolaus Copernicus University, Toruń, Poland
e-mail: : wniemiro@gmail.com

Ryszard Zieliński

Inst. Math. Polish Acad. Sc.
e-mail: R.Zielinski@impan.gov.pl

Abstract. For every probability of success $\theta \in]0,1[$, the sequence of Bernoulli trials is asymptotically normal, but it is *not uniformly* in $\theta \in]0,1[$ normal. We show that the uniform asymptotic normality holds if the sequence of Bernoulli trials is randomly stopped with an appropriate stopping rule.

Mathematics Subject Classification: 60F05, 60B10, 62L12

Key words and phrases: Central Limit Theorem, Uniform Central Limit Theorem, Bernoulli scheme, stopping times, sequential estimators

1. Introduction

For the Bernoulli scheme with a probability of success θ, the central limit theorem (CLT) does *not* hold uniformly in $\theta \in]0,1[$: for any fixed n (the number of trials), the normal approximation fails and its error is close to $1/2$ if θ is close to 0 (Zieliński 2004). CLT does not hold also for the negative Bernoulli scheme (ibid.). In our paper we show that CLT holds if n is an appropriate random variable. A sequence of stopping times and estimators are effectively constructed.

2. Main Results

Let $Z_1, \ldots, Z_n, \ldots$ be a sequence of random variables defined on a statistical space with a family of distributions $\{P_\theta : \theta \in \Theta\}$.

2.1. Definition. *The sequence Z_n is uniformly asymptotically normal (UAN) if for some functions $\mu(\theta)$ and $\sigma^2(\theta)$,*

$$\forall_\varepsilon \exists_{n_0} \forall_{n \geq n_0} \forall_\theta \sup_{-\infty < x < \infty} \left| P_\theta\left(\frac{\sqrt{n}}{\sigma(\theta)}[Z_n - \mu(\theta)] \leq x\right) - \Phi(x)\right| < \varepsilon,$$

where Φ is the c.d.f. of the standard normal distribution $N(0,1)$. We will then write

$$\frac{\sqrt{n}}{\sigma(\theta)}[Z_n - \mu(\theta)] \rightrightarrows N(0,1).$$

Uniform convergence in distribution is considered e.g. in Zieliński 2004, Salibian-Barrera and Zamar (2004), and Borovkov (1998). The definition above may be considered as a special case of that in Borovkov 1998.

2.2. Theorem.*Let $X = X_1, \ldots, X_n, \ldots$ be i.i.d. with $P_\theta(X = 1) = \theta = 1 - P_\theta(X = 0)$. The parameter space is $\Theta =]0,1[$.*

(i) There is no sequence of estimators $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ such that

$$\frac{\sqrt{n}}{\sigma(\theta)}[\hat{\theta}_n - \theta] \rightrightarrows N(0,1).$$

(ii) There is a sequence of stopping rules T_r $(r = 1, 2, \ldots)$ and a sequence of estimators $\hat{\theta}_r = \hat{\theta}_r(X_1, \ldots, X_{T_r})$ such that

$$\frac{\sqrt{r}}{\sigma(\theta)}[\hat{\theta}_r - \theta] \rightrightarrows N(0,1).$$

Proof of part (i). For every n there exists θ such that $P_\theta(X_1 = \cdots = X_n = 0) > 1/2$. For such θ the probability distribution of the random variable $(\sqrt{n}/\sigma(\theta))[\hat{\theta}_n - \theta]$ has an atom which contains more than $1/2$ of the total probability mass. It follows that

$$\sup_{-\infty<x<\infty} \left| P_\theta[(\sqrt{n}/\sigma(\theta))[\hat{\theta}_n - \theta] \le x] - \Phi(x) \right| \ge 1/4.$$

□

The proof of part (ii) requires some auxiliary lemmas and will be presented in details in next sections.

3. Proofs

3.1. Lemma (A uniform version of the δ-method). *Let h be a function differentiable at μ. Assume that h and μ do not depend on θ. If*

$$V_n = \frac{\sqrt{n}}{\sigma(\theta)}[Z_n - \mu] \rightrightarrows N(0,1),$$

$h'(\mu) \neq 0$ and $\sigma(\theta) \le b$ for some $b < \infty$ and for all $\theta \in (0,1)$ then

$$\frac{\sqrt{n}}{\sigma(\theta)h'(\mu)}[h(Z_n) - h(\mu)] \rightrightarrows N(0,1).$$

Proof. Obviously $h(z) - h(\mu) = h'(\mu)(z-\mu) + r(z)(z-\mu)$, where $r(z) \to 0$ as $z \to \mu$, and in consequence

$$\frac{\sqrt{n}}{\sigma(\theta)h'(\mu)}[h(Z_n) - h(\mu)] = V_n + R_n$$

where

$$R_n = \frac{r(Z_n)}{h'(\mu)} \frac{\sqrt{n}}{\sigma(\theta)} [Z_n - \mu].$$

We will show that R_n tends to zero uniformly in probability P_θ, i.e. that for every $\delta > 0$,

$$\textbf{(3.2)} \qquad \sup_{0<\theta<1} P_\theta(|R_n| > \delta) \to 0.$$

To this end fix $\delta > 0$ and $\varepsilon > 0$ and choose a such that $1-\Phi(a)+\Phi(-a) < \varepsilon$. For sufficiently large n we have

$$\sup_{|z-\mu|\le ab/\sqrt{n}} \left| \frac{r(z)}{h'(\mu)} \right| < \frac{\delta}{a}.$$

If the inequality holds then on the event $\{|V_n| \le a\}$ we have $|Z_n - \mu| = |V_n|\sigma(\theta)/\sqrt{n} \le ab/\sqrt{n}$ and consequently $|R_n| = |r(Z_n)/h'(\mu)| \cdot |V_n| < \delta$. For sufficiently large n we also have $\sup_\theta \sup_x |P_\theta(V_n \le x) - \Phi(x)| < \varepsilon$ and therefore

$$\begin{aligned} \sup_\theta P_\theta(|R_n| > \delta) &\le \sup_\theta P_\theta(|V_n| > a) \\ &\le 1 - \Phi(a) + \Phi(-a) + 2\varepsilon < 3\varepsilon, \end{aligned}$$

which ends the proof of (3.2). We end the proof of Lemma 3.1 using the following inequalities

$$\begin{aligned} P_\theta\left(V_n + R_n \le x\right) &\le P_\theta\left(V_n \le x + \delta\right) + P_\theta\left(|R_n| > \delta\right), \\ P_\theta\left(V_n + R_n \le x\right) &\ge P_\theta\left(V_n \le x - \delta\right) - P_\theta\left(|R_n| > \delta\right), \end{aligned}$$

and the uniform continuity of Φ. □

3.3. Berry-Esséen Theorem. By the standard Berry-Esséen Theorem for *i.i.d.* random variables $Y_1, \ldots, Y_n, \ldots$, $S_n = \sum_1^n Y_i$, and $F_n(x) = P(n^{-1/2}\sigma^{-1}[S_n - n\mu] \le x)$ we have

$$|F_n(x) - \Phi(x)| \le C \frac{m_3}{\sigma^3 \sqrt{n}},$$

where $m_3 = E|Y - \mu|^3$ and C is an absolute constant.

By the following sequence of inequalities $m_3^{1/3} \le m_4^{1/4}$, $\sigma = m_2^{1/2} \le m_4^{1/4}$, and

$$\frac{m_3}{\sigma^3} \le \frac{m_4^{3/4}}{\sigma^3} = \frac{m_4^{3/4}}{\sigma^4}\sigma \le \frac{m_4^{3/4}}{\sigma^4} m_4^{1/4} = \frac{m_4}{\sigma^4}$$

we obtain

3.4. Corollary

$$|F_n(x) - \Phi(x)| \le C \frac{m_4}{\sigma^4 \sqrt{n}},$$

where $m_4 = E(Y - \mu)^4$.

Let us now consider the negative binomial scheme, that is an i.i.d. sequence of random variables geometrically distributed with the parameter θ. The central limit theorem for this scheme does not hold uniformly in $\theta \in]0,1[$ (Zieliński 2004): the normal approximation breaks down for θ approaching 1. In the following lemma we assume θ to be bounded away from 1.

3.5. Lemma [Central Limit Theorem for the negative binomial scheme]. *Let* $Y = Y_1, \ldots, Y_r, \ldots$ *be i.i.d. and let* $P_\theta(Y = k) = \theta(1-\theta)^{k-1}$ *for* $k = 1, 2, \ldots$. *Let* $T_r = \sum_1^r Y_i$. *Assume that* $\theta \le 1 - \kappa$: *the parameter space is* $\Theta =]0, 1-\kappa]$ *for some* $\kappa > 0$. *Then*

$$\frac{\sqrt{r}}{\sqrt{1-\theta}} \left(\frac{\theta T_r}{r} - 1 \right) \rightrightarrows N(0,1).$$

We will use following elementary facts about the geometric distribution

$$E_\theta(Y) = \frac{1}{\theta}, \qquad \sigma^2(\theta) = Var_\theta(Y) = \frac{1-\theta}{\theta^2},$$

and

$$m_4(\theta) = E_\theta(Y - \mu(\theta))^4 = \frac{(1-\theta)(\theta^2 - 9\theta + 9)}{\theta^4}.$$

Consequently, for $\theta \le 1 - \kappa$,

$$\frac{m_4(\theta)}{\sigma^4(\theta)} = \frac{\theta^2 - 9\theta + 9}{1-\theta} = \frac{\theta^2}{1-\theta} + 9 \le \frac{1}{\kappa} + 9.$$

From Corollary 3.4 it follows that

$$\sqrt{r}\frac{\theta}{\sqrt{1-\theta}}\left(\frac{T_r}{r}-\frac{1}{\theta}\right)\rightrightarrows N(0,1) \quad \text{uniformly in } \theta\in]1,1-\kappa].$$

□

3.6. Lemma. *Under the assumptions of the previous lemma,*

$$\frac{\sqrt{r}}{\sqrt{1-\theta}}\left(\frac{r}{\theta T_r}-1\right)\rightrightarrows N(0,1).$$

Proof. It is enough to combine Lemma 3.6 with Lemma 3.1 (δ-method) applied to the function $h(x)=1/x$ at $\mu=1$. □

3.7. Lemma. *Let $X_1,\ldots,X_n,\ldots$ be the Bernoulli scheme with a probability of success θ. Define the sequence of stopping rules $T'_r=\min\{n: S_n\geq r\}$, where $S_n=\sum_1^n X_i$. The sequence $\hat{\theta}'_r=r/T'_r$ is UAN in $\theta\leq 1-\kappa$, i.e. for the parameter space $\Theta=]0,1-\kappa]$.*

Proof. This is a simple reformulation of Lemma 3.6. Indeed, it is easy to see that T'_r is a sum of i.i.d.geometrically distributed random variables.

Proof of Theorem 2.2(ii). The sequence of stopping times $T_r, r=1,2,\ldots,$ will be constructed as follows. Define $T'_r=\min\{n: S_n\geq r\}$, $T''_r=\min\{n: n-S_n\geq r\}$,

$$\tilde{T}_r=\min\{n: S_n\geq r, n-S_n\geq r\}=\max(T'_r,T''_r),$$

and

$$T_r=\tilde{T}_r+r.$$

The sequence of estimators $\hat{\theta}_r$ will be constructed as follows. Define two auxiliary estimators $\hat{\theta}'_r=r/T'_r$ and $\hat{\theta}''_r=1-r/T''_r$, a random event

$$A_r=\left\{\frac{1}{r}\sum_{i=1}^{r}X_{\tilde{T}_r+i}<\frac{1}{2}\right\},$$

and finally

$$\hat{\theta}_r = \begin{cases} \hat{\theta}'_r & \text{on } A_r \\ \hat{\theta}''_r & \text{on } A_r^c. \end{cases}$$

We claim that $\hat{\theta}_r$ is UAN on $]0,1[$ with the asymptotic variance $\sigma^2(\theta)$ given by the formula:

$$\sigma^2(\theta) = \begin{cases} (1-\theta)\theta^2 & \text{for } \theta < 1/2, \\ (1-\theta)^2\theta & \text{for } \theta \geq 1/2. \end{cases}$$

To prove that fix $\varepsilon > 0$ and choose $\delta > 0$ such that

$$\sup_{1/2-\delta<\theta<1/2+\delta} \sup_x \left| \Phi\left(\frac{x}{\theta\sqrt{1-\theta}}\right) - \Phi\left(\frac{x}{\sqrt{\theta}(1-\theta)}\right) \right| < \varepsilon.$$

Obviously $\delta < 1/2$.

Choose r_1 such that for $r \geq r_1$ the inequality $P_\theta(A_r^c) < \varepsilon$ holds for all $\theta < 1/2 - \delta$ and $P_\theta(A_r) < \varepsilon$ holds for all $\theta > 1/2 + \delta$.

From Lemma 3.7 we conclude that

$$\frac{\sqrt{r}}{\theta\sqrt{1-\theta}}\left(\hat{\theta}'_r - \theta\right) \rightrightarrows N(0,1) \quad \text{on} \quad]0, 1/2+\delta]$$

and

$$\frac{\sqrt{r}}{\sqrt{\theta}(1-\theta)}\left(\hat{\theta}''_r - \theta\right) \rightrightarrows N(0,1) \quad \text{on} \quad [1/2-\delta, 1[.$$

Choose r_2 such that for $r \geq r_2$ and for all $\theta \leq 1/2 + \delta$,

$$\sup_x \left| P_\theta\left(\sqrt{r}\frac{\hat{\theta}'_r - \theta}{\theta\sqrt{1-\theta}} \leq x\right) - \Phi(x) \right| \\ = \sup_x \left| P_\theta\left(\sqrt{r}(\hat{\theta}'_r - \theta) \leq x\right) - \Phi\left(\frac{x}{\theta\sqrt{1-\theta}}\right) \right| < \varepsilon.$$

Then for $r \geq r_2$ and for all $\theta \geq 1/2 - \delta$ we also have

$$\sup_x \left| P_\theta\left(\sqrt{r}\frac{\hat{\theta}''_r - \theta}{\sqrt{\theta}(1-\theta)} \leq x\right) - \Phi(x) \right| \\ = \sup_x \left| P_\theta\left(\sqrt{r}(\hat{\theta}''_r - \theta) \leq x\right) - \Phi\left(\frac{x}{\sqrt{\theta}(1-\theta)}\right) \right| < \varepsilon.$$

Define $r_0 = \max(r_1, r_2)$.

For the estimator $\hat{\theta}_r$ we obtain

$$\begin{aligned}
&\sup_x \left| P_\theta\left(\sqrt{r}(\hat{\theta}_r - \theta) \le x\right) - \Phi\left(\frac{x}{\sigma(\theta)}\right) \right| \\
&\quad \le \sup_x \left| P_\theta\left(\sqrt{r}(\hat{\theta}_r - \theta) \le x, A_r\right) - P_\theta(A_r)\Phi\left(\frac{x}{\sigma(\theta)}\right) \right| \\
&\qquad + \sup_x \left| P_\theta\left(\sqrt{r}(\hat{\theta}_r - \theta) \le x, A_r^c\right) - P_\theta(A_r^c)\Phi\left(\frac{x}{\sigma(\theta)}\right) \right|.
\end{aligned}$$

Due to the facts that $\hat{\theta}_r = \hat{\theta}'_r$ on A_r and $\hat{\theta}'_r$ and A_r are independent, and similarly $\hat{\theta}_r = \hat{\theta}''_r$ on A_r^c and $\hat{\theta}''_r$ and A_r^c are independent, the Right Hand Side of the latter formula is equal to

$$\begin{aligned}
&P_\theta(A_r) \cdot \sup_x \left| P_\theta\left(\sqrt{r}(\hat{\theta}'_r - \theta) \le x\right) - \Phi\left(\frac{x}{\sigma(\theta)}\right) \right| \\
&\quad + P_\theta(A_r^c) \cdot \sup_x \left| P_\theta\left(\sqrt{r}(\hat{\theta}''_r - \theta) \le x\right) - \Phi\left(\frac{x}{\sigma(\theta)}\right) \right|.
\end{aligned}$$

For $\theta < 1/2 - \delta < 1/2$ we have $P_\theta(A_r^c) < \varepsilon$, $\sigma^2(\theta) = (1-\theta)\theta^2$, and

$$\left| P_\theta\left(\sqrt{r}(\hat{\theta}'_r - \theta) \le x\right) - \Phi\left(\frac{x}{\theta\sqrt{1-\theta}}\right) \right| < \varepsilon.$$

For $\theta > 1/2 + \delta > 1/2$ we have $P_\theta(A_r) < \varepsilon$, $\sigma^2(\theta) = (1-\theta)^2\theta$, and

$$\left| P_\theta\left(\sqrt{r}(\hat{\theta}''_r - \theta) \le x\right) - \Phi\left(\frac{x}{\sqrt{\theta}(1-\theta)}\right) \right| < \varepsilon.$$

For $1/2 - \delta < \theta < 1/2 + \delta$

$$\begin{aligned}
&\left| P_\theta\left(\sqrt{r}(\hat{\theta}'_r - \theta) \le x\right) - \Phi\left(\frac{x}{\sigma(\theta)}\right) \right| \\
&< \left| P_\theta\left(\sqrt{r}(\hat{\theta}'_r - \theta) \le x\right) - \Phi\left(\frac{x}{\theta\sqrt{1-\theta}}\right) \right| + \left| \Phi\left(\frac{x}{\theta\sqrt{1-\theta}}\right) - \Phi\left(\frac{x}{\sigma(\theta)}\right) \right| \\
&< 2\varepsilon
\end{aligned}$$

and similarly

$$
\begin{aligned}
&\left|P_\theta\left(\sqrt{r}(\hat{\theta}_r'' - \theta) \leq x\right) - \Phi\left(\frac{x}{\sigma(\theta)}\right)\right| \\
&< \left|P_\theta\left(\sqrt{r}(\hat{\theta}_r'' - \theta) \leq x\right) - \Phi\left(\frac{x}{\sqrt{\theta}(1-\theta)}\right)\right| + \left|\Phi\left(\frac{x}{\sqrt{\theta}(1-\theta)}\right) - \Phi\left(\frac{x}{\sigma(\theta)}\right)\right| \\
&< 2\varepsilon.
\end{aligned}
$$

Eventually we obtain

$$
\sup_x \left|P_\theta\left(\sqrt{r}(\hat{\theta}_r - \theta) \leq x\right) - \Phi\left(\frac{x}{\sigma(\theta)}\right)\right| < 4\varepsilon
$$

which ends the proof. □

References

Borovkov, A. A. (1998). *Mathematical Statistics*, Gordon and Breach.

Feller, W. (1966). *An Introduction to Probability Theory and its Applications*, Vol. II, Wiley.

Salibian-Barrera, M. and Zamar, R. H. (2004). Uniform asymptotics for robust location estimates when the scale is unknown, *Ann. Statist.* 32, 4, 1434–1447.

Zieliński R. (2004). Effective WLLN, SLLN and CLT in statistical models, *Applicationes Mathemticae* 31, 1, 117–125