On the T-test

S.Y.Novak*

MDX University London

Abstract

The aim of this article is to show that the T-test can be misleading. We argue that normal or Student's approximation to the distribution $\mathcal{L}(t_n)$ of Student's statistic t_n does not hold uniformly over the class $\mathcal{P}_n$ of samples $\{X_1, ..., X_n\}$ from zero-mean unit-variance bounded distributions. We present lower bounds to the corresponding error.

We suggest a generalisation of the T-test that allows for variability of possible approximating distributions to $\mathcal{L}(t_n)$.

Key words: Hypothesis testing, T-test, Student's statistic.
AMS Subject Classification: 60E15, 62G10, 62G35.

Given a sample $X_1, ..., X_n$ of independent and identically distributed (i.i.d.) observations over a random variable (r.v.) X, denote

$$t_n = (\hat{X} - \mathbb{E}X)\sqrt{n}/\hat{\sigma} ,$$

where $\hat{X} = S_n/n$, $S_n = X_1 + ... + X_n$, and $\hat{\sigma}$ is an estimator of the standard deviation of X. In hypothesis testing the test of the hypothesis $H_0 = \{\mathbb{E}X = a\}$ involving test statistic t_n is called the T-test; r.v. t_n is Student's statistic.

T-test is one of the most widely used statistical tests. Textbooks advocate using the T-test when testing hypothesis H_0 vs the alternative hypothesis $H_A = \{\mathbb{E}X = b\}$, where $a \neq b$; when testing hypothesis $\{\mathbb{E}X \leq a\}$ vs hypothesis $\{\mathbb{E}X \geq b\}$, etc..

In view of the law of large numbers and the central limit theorem the T-test appears perfectly justified if $\mathbb{E}X^2 < \infty$ and the sample size is large: "the size of the one- and two-sample T-tests is relatively insensitive to nonnormality (at least for large samples). Power values of the T-tests obtained under normality are asymptotically valid also for all other distributions with finite variance." ([3], p. 207).

We show below that the T-test has problems even in the simplest situation where $\sigma^2 := \mathrm{var}\, X$ is known. We argue that the T-test is not automatically applicable, and requires prior checks.

The reason for that is that the test is effectively applied as a non-parametric one — textbooks implicitly assume that the T-test "works" uniformly over the *non-parametric* class $\mathcal{P}_\sigma(a_1, a_2)$ of distributions with mean $\mathbb{E}X \in [a_1; a_2]$ and standard deviation σ.

We show that weak convergence of $(S_n - \mathbb{E}S_n)/\sqrt{n}$ to the normal law cannot hold uniformly in the class of zero-mean unit-variance distributions (the issue with uniform convergence is known

*Supported by the Engineering and Physical Research Council grant No EP/W010607/1.

in the literature though not in the context of the T-test — see, e.g., [6] and references therein). In particular, normal or Student's approximation to the distribution of Student's statistic is not automatically applicable.

We suggest performing prior checks in order to find out if a particular (not necessarily normal or Student's) approximation to the distribution of the test statistic is applicable. This leads to a generalisation of the T-test that allows for non-conventional approximating distributions. We discuss implications for the choice of critical levels.

Section 1 addresses the question if the T-test is applicable uniformly over class $\mathcal{P}_n$. Section 2 presents an example of non-normal approximation to $\mathcal{L}(t_n)$ as well as an estimate of the accuracy of such approximation in terms of the total variation distance. The approximating distribution appears new in the literature on the topic. Section 3 suggests a generalisation of the T-test. Proofs are postponed to section 4.

1 Problems with the T-test

The T-test has been criticized by a number of authors. For instance, Bahadur ([2], Example 8.1) shows that the T-test is not Bahadur-efficient if $H_0 = \{\mathbb{E}X = 0\}$ and $X_1, ..., X_n$ are i.i.d. normal $\mathcal{N}(\theta; 1)$ r.v., where $\theta \geq 0$. Rukhin [12] shows that the T-test is not Bahadur-efficient in the case of testing the null hypothesis $H_0 = \{\theta = 0\}$ against $H_A = \{\theta = b\}$ for the parametric family $\{F_{\theta,c}, \theta \in \mathbb{R}, c > 0\}$, where $F_{\theta,c}(x) = F((x-\theta)/c)$ $(\forall x)$, F is a distribution function (d.f.) with a finite (in a neighbourhood of 0) moment generating function.

The T-test is usually applied in the assumption that the underlying distribution has a finite variance. We show below that the use of the T-test is not justified even in the case of testing a simple hypothesis $H_0 = \{\mathbb{E}X = a\}$ against a simple alternative $H_A = \{\mathbb{E}X = b\}$ in the assumption that $\mathrm{var}\, X < \infty$. W.l.o.g. we may assume in the sequel that $a = 0$, i.e., $H_0 = \{\mathbb{E}X = 0\}$.

Let $\mathcal{P}_n$ denote the class of distributions $\mathcal{L}(X_1, ..., X_n)$ such that $X, X_1, ..., X_n$ are i.i.d. bounded r.v.s, $\mathbb{E}X = 0$, $\mathbb{E}X^2 = 1$. The use of normal or Student's approximation in the T-test would be justified if such approximation held uniformly in class $\mathcal{P}_n$.

We show below that normal and Student's approximation to $\mathcal{L}(t_n)$ does not hold uniformly in the class $\mathcal{P}_n$. Namely, there exists an absolute constant $c > 0$ such that for any $n > 12$

$$\inf_{x \geq 0} \sup_{\mathcal{P}_n} |\mathbb{P}(t_n \geq x)/\Phi_c(x) - 1| \geq c, \tag{1}$$

where Φ denotes the standard normal distribution function, $\Phi_c = 1 - \Phi$.

A similar result holds if standard normal d.f. Φ in (1) is replaced with F_n or F_{n-k}, where F_n denotes the distribution function of Student's statistic with n degrees of freedom, $k \in \mathbb{N}$. Thus, the T-test is not applicable uniformly over $\mathcal{P}_n$; the outcome of the test can be misleading even for large-size samples.

Note that F_n is close to Φ:

$$\sup_x |F_n(x) - \Phi(x)| \leq C/n \qquad (n \to \infty) \tag{2}$$

(cf. Pinelis [10]). The table of Student's distribution function shows little difference between $F_n(\cdot)$ and $\Phi(\cdot)$ if $n \geq 60$. Thus, preference to F_n over Φ appears questionable.

Theorem 1 *As* $n\to\infty$,

$$\inf_{x\geq 0}\sup_{\mathcal{P}_n}\left|\mathbb{P}(t_n^*\geq x)/\Phi_c(x)-1\right| \geq 1/4+O(1/n). \tag{3}$$

If $\{x_n\}$ *is a non-decreasing sequence of positive numbers such that* $1\ll x_n\leq\sqrt{n}$ *as* $n\to\infty$, *then*

$$\sup_{\mathcal{P}_n}\left|\mathbb{P}(t_n\geq x_n)/\Phi_c(x_n)-1\right| \to\infty \qquad (n\to\infty). \tag{4}$$

A similar result holds if normal approximation to $\mathcal{L}(t_n)$ has been replaced with Student's approximation. Denote $F_n^c=1-F_n$.

Theorem 2 *As* $n\to\infty$,

$$\inf_{x\geq 0}\sup_{\mathcal{P}_n}\left|\mathbb{P}(t_n\geq x)/F_n^c(x)-1\right| \geq 1/4+O(1/n). \tag{5}$$

If $\{x_n\}$ *is a non-decreasing sequence of positive numbers such that* $1\ll x_n\leq\sqrt{n}$ *as* $n\to\infty$, *then*

$$\sup_{\mathcal{P}_n}\left|\mathbb{P}(t_n\geq x_n)/F_n^c(x_n)-1\right| \to\infty \qquad (n\to\infty). \tag{5*}$$

The result holds if F_n *in (5) has been replaced with* F_{n-k}, *where* k *is a fixed natural number.*

Note that critical values of the T-test are determined by the limiting distribution of t_n, probabilities of the type-II error are large deviations probabilities like $\mathbb{P}(t_n\geq c\sqrt{n})$ (see, e.g., [8]). Theorems 1, 2 show that the probabilities of type-I and type-II errors in the T-test can be very different from those traditionally assumed.

2 An example of non-normal approximation

It may be counter-intuitive to expect that Poisson distribution may play any role in relation to the T-test. However, Proposition 3 below states it may.

In this section we present an example of non-normal/non-Student's approximation to the distribution of Student's statistic t_n and the self-normalised sum

$$t_n^*=S_n/T_n^{1/2},$$

where $T_n=\sum_{i=1}^n X_i^2$. We evaluate the accuracy of such approximation.

Self-normalised sum t_n^* is closely related to Student's statistic t_n:

$$t_n=t_n^*/\sqrt{1-t_n^{*\,2}/n}\,,\; t_n^*=t_n/\sqrt{1+t_n^2/n}\,. \tag{6}$$

Therefore,

$$\{t_n\geq x\}=\left\{\, t_n^*\geq x/\sqrt{1+x^2/n}\,\right\},\; \{t_n^*\geq y\}=\left\{\, t_n\geq y/\sqrt{1-y^2/n}\,\right\} \tag{6*}$$

if $x\geq 0,\ 0\leq y\leq\sqrt{n}$. Thus, the limiting distributions of t_n and t_n^* coincide.

The example below highlights the fact that $\mathcal{L}(t_n)$ as well as the limiting distribution of Student's statistic may take on value ∞ with positive probability.

Given r.v.s ξ and η, we denote by $d_{TV}(\xi;\eta) \equiv d_{TV}(\mathcal{L}(\xi);\mathcal{L}(\eta))$ the total variation distance between $\mathcal{L}(\xi)$ and $\mathcal{L}(\eta)$. Let π_λ denote a Poisson r.v. with parameter λ. Set

$$Y_n = (np-\pi_{np})\Big/\sqrt{\pi_{np}(1-\pi_{np}/n)}\,, \quad Y_n^* = (np-\pi_{np})\Big/\sqrt{np^2+(1-2p)\pi_{np}}\,, \tag{7}$$

where $p\in(0;1/2]$. Note that

$$\mathbb{P}(Y_n=\sqrt{n}) = e^{-np}\,.$$

Proposition 3 *Let $X, X_1, ..., X_n$ be i.i.d.r.v.s with the distribution*

$$\mathbb{P}\left(X=\sqrt{p/q}\right) = q,\ \mathbb{P}\left(X=-\sqrt{q/p}\right) = p, \tag{8}$$

where $p\in(0;1/4]$, $q=1-p$. Then

$$d_{TV}(t_n;Y_n) \le 3p/4e + 4(1-e^{-np})p^2\,. \tag{9}$$

In the light of (6), inequality (9) can be reformulated as follows:

$$d_{TV}(t_n^*;Y_n^*) \le 3p/4e + 4(1-e^{-np})p^2\,. \tag{9^+}$$

Given $\lambda>0$, denote

$$Y(\lambda) = (\lambda-\pi_\lambda)/\sqrt{\pi_\lambda}\,.$$

Clearly, $Y(\lambda)$ is a defective random variable: $Y(\lambda)$ takes on value ∞ with probability $e^{-\lambda}$. According to Proposition 3,

$$t_n \Rightarrow Y(\lambda),\ t_n^* \Rightarrow Y(\lambda) \qquad (n\to\infty) \tag{10}$$

if $p=p(n)\sim\lambda/n$ as $n\to\infty$.

Weak convergence (10) may hold in more general situations, e.g., if $X_i \stackrel{d}{=} (\xi_i-\mathbb{E}\xi)/\mathbb{E}^{1/2}\xi$ and $\xi,\xi_1,\xi_2,...,\xi_n$ are i.i.d. non-degenerate r.v.s taking values in $\mathbb{Z}_+=\{0,1,2,...\}$. For example, (10) holds if $X_i \stackrel{d}{=} (p-\eta_i)/\sqrt{p}$, where $\{\eta_i\}$ are i.i.d. Poisson $\mathbf{\Pi}(p)$ r.v.s with $p=p(n)\sim\lambda/n$ as $n\to\infty$.

In situations where t_n can be approximated by Y_n or Z_n the "asymptotic approach" suggests the critical values $c_- \equiv c_-(\varepsilon)$ and $c_+ \equiv c_+(\varepsilon)$ of the two-sided T-test be chosen according to equations

$$\mathbb{P}(Y(\lambda)>c_+) = \mathbb{P}(Y(\lambda)<c_-) = \varepsilon/2 \qquad (\varepsilon>0)$$

with $\lambda=np$ replaced by its consistent estimator; the "sub-asymptotic approach" (cf. [5], ch. 9) suggests incorporating estimate (9).

3 A generalised test

The T-test relies on the validity of normal (or Student's) approximation to $\mathcal{L}(t_n)$. The common impression is that $\mathcal{L}(t_n)$ is close to the standard normal distribution if the sample size n is large (see, e.g., Lehman [3], p. 205). However, it is known that the limiting distribution of t_n is not always normal (the class $\mathcal{L}_\mathcal{S}$ of limiting distributions of Student's statistic has been described by Mason [4]).

In this section we suggest a generalised T-test. The idea is to check first if a particular approximation (not necessarily normal or Student's) is applicable. The latter can be done using estimates of the accuracy of approximation.

Thus, the generalised T-test requires
(1) a list of possible limiting/approximating distributions;
(2) sharp estimates of the accuracy of approximation of $\mathcal{L}(t_n)$ by a particular distribution;
(3) estimation of certain quantities involved in those estimates of the accuracy of approximation (e.g., estimation of σ and $\mathbb{E}|X^3|$ in the case of normal approximation).

Traditionally, the obvious candidate for the approximating distribution is the standard normal law $\mathcal{N}(0;1)$. One can employ the following approximate bound to the uniform distance between $\mathcal{L}(t_n^*)$ and $\mathcal{N}(0;1)$ (cf. [5], Corollary 12.22): for all large enough n

$$|\mathbb{P}\,(t_n\!<\!x) - \Phi(x)| \le (6.4\hat{\mu}_3/\hat{\sigma}^3 + 2\hat{\mu}_1/\hat{\sigma})/\sqrt{n}\,, \tag{11}$$

where $\hat{\mu}_k$ denotes a consistent estimator of $\mu_k := \mathbb{E}|X-\mathbb{E}X|^k$ $(k\ge 1)$; $\hat{\sigma}$ is an estimator of the standard deviation of X.

Bound (11) seems to be the sharpest available in the case of i.i.d. observations (cf. the discussion in [11], Remarks 4.16–4.17).

The use of normal approximation can be justified if the right-hand side (r.h.s.) of (11) is less than a certain small number (say, ε_o) specified by a statistician (e.g., $\varepsilon_o\!=\!0.01$).

Since the limiting distribution of t_n may differ from $\mathcal{N}(0;1)$, we suggest that one first checks if a particular (not necessarily normal) approximation to $\mathcal{L}(t_n)$ is applicable.

One may have a number of bounds of the type

$$\sup_x |\mathbb{P}(t_n\!\le\! x) - G_k(x)| \le r_n(k), \tag{12}$$

where $G_1, G_2, ...$ are d.f.s of certain candidate distributions. It is natural to choose $k=k_*$ such that $r_n(k_*) = \min_k r_n(k)$. Note that for most distributions from $\mathcal{L}_{\mathcal{S}}$ the task of deriving estimates of the accuracy of approximation with explicit constants remains open.

Obviously, one needs a list of possible approximating distributions together with the corresponding estimates of the accuracy of approximation (with explicit constants). Such a list will always be finite (until recently only normal and Student's distributions were on the list). Proposition 3 adds another candidate to that list.

The problem of deriving estimates of the accuracy of normal approximation with explicit constants to the distribution of a sum of r.v.s goes back to Tchebychef [14]. It led to a vast literature with contributions from many renowned authors (see, e.g., references in [1, 5, 9, 13]). The task of evaluating the accuracy of Poisson and compound Poisson approximation has been addressed by many distinguished authors (see, e.g., references in [1, 7]).

Note that one can have a situation where neither distribution from the list has the estimate $r_n(k)$ of the accuracy of approximation below the specified threshold level ε_o (i.e., $\min_k r_n(k)\!>\!\varepsilon_o$). That would mean the T-test is not applicable (either because of a small sample size or because of the list being too short).

4 Proofs

Since t_n and t_n^* are scale-invariant, w.l.o.g. we may assume in the sequel that $\operatorname{var} X = 1$. The proofs of Theorems 1, 2 use the fact that $\mathcal{L}(t_n)$ and $\mathcal{L}(t_n^*)$ are not stochastically bounded uniformly in $\mathcal{P}_n$. Below the operation of multiplication is superior to the division.

Proof of Theorem 1. Taking into account (6*), we shall show that

$$\inf_{x\geq 0}\sup_{\mathcal{P}_n}\Big|\mathbb{P}(t_n^*\geq x)/\Phi_c(x)-1\Big| \geq 1.25e^{-1/2(n-2)}-1>0 \tag{13}$$

as $n>12$.

Note that $t_n^*\leq\sqrt{n}$. Thus, (13) trivially holds if $x>\sqrt{n}$. Therefore, we may assume in the sequel that $x\in[0;\sqrt{n}\,]$.

It suffices finding i.i.d. bounded r.v.s $X,X_1,...,X_n$ such that $\mathbb{E}X=0$, $\mathbb{E}X^2=1$, and (13) holds. We employ distribution (8) that seems to play the role of a testing stone when one deals with self-normalised sums and Student's statistic (cf. Example 12.3 in [5]).

Let X be a r.v. with distribution (8), where $p\in(0;1/4]$, $q=1-p$. Then

$$X_i \stackrel{d}{=} (p-\xi_i)/\sqrt{pq} \qquad (i\geq 1), \tag{8*}$$

where $\{\xi_i\}$ are independent Bernoulli $\mathbf{B}(p)$ r.v.s. Note that

$$\mathbb{E}X=0,\ \mathbb{E}X^2=1,\ \mathbb{E}|X|^3=(p^2+q^2)/\sqrt{pq}\,.$$

Hence $\mathcal{L}(X_1,...,X_n)\in\mathcal{P}_n$.

Denote $S_n^\xi=\xi_1+...+\xi_n$. Then

$$\begin{aligned} S_n &= (np-S_n^\xi)/\sqrt{pq}\,,\quad T_n=np/q+(1-2p)S_n^\xi/pq,\\ t_n^* &= (np-S_n^\xi)/\sqrt{np^2+(q-p)S_n^\xi}\,. \end{aligned} \tag{14}$$

Set

$$g(k)=(np-k)/\sqrt{np^2+(q-p)k} \qquad (k\in\mathbb{Z}_+). \tag{15}$$

Note that $t_n^*=g(S_n^\xi)$. Since function $g(\cdot)\downarrow$, we have

$$\mathbb{P}(t_n^*\geq g(k))=\mathbb{P}(S_n^\xi\leq k). \tag{16}$$

Clearly, t_n^* takes on its largest possible value $g(0)=\sqrt{n}$ when $X_1=...=X_n=\sqrt{p/q}$, t_n^* takes on its second largest possible value $g(1)=(np-1)/\sqrt{np^2+q-p}$ when $n-1$ sample elements equal $\sqrt{p/q}$ and one sample element equals $-\sqrt{q/p}$, etc.. Hence

$$\mathbb{P}(t_n^*=\sqrt{n}\,)=q^n,\ \mathbb{P}\Big(t_n^*=(np-1)/\sqrt{np^2+(q-p)}\Big)=npq^{n-1}. \tag{17}$$

We consider first the case where $x\in[0;1]$. According to (16), (17),

$$\mathbb{P}(t_n^*\geq g(1))=(q+np)q^{n-1}.$$

Note that

$$\ln(1-x)\geq -x-x^2/2(1-x)^2 \qquad (0\leq x<1).$$

Hence

$$(1-p)^n\geq\exp(-np(1+p/2q^2)). \tag{18}$$

Denote
$$p_x = \Big(1+x\sqrt{1-1/n}\Big/\sqrt{1-x^2/n}\,\Big)\Big/n\,.$$
Set $p=p_x$. Then $g(1)=x$.

One can check that $np/q\geq 1+x$. Hence
$$\mathbb{P}(t_n^*\geq x)\geq (2+x)q^n.$$
Taking into account (18), we derive
$$\begin{aligned}\mathbb{P}(t_n^*\geq x) &\geq (2+x)\exp\Big(-\Big(1+x\sqrt{1-1/n}\Big/\sqrt{1-x^2/n}\,\Big)(1+p/2q^2)\Big)\\ &\geq (2+x)\exp\Big(-(1+x)\Big(1+(1+x)/2n(1-2/n)^2\Big)\Big).\end{aligned}$$

Denote
$$f(x)=\tfrac{2}{e}(2+x)\exp(x^2/2-x-2/n(1-2/n)^2.$$
It is well-known that $\Phi_c(x)\leq\frac12 e^{-x^2/2}$. Hence
$$\mathbb{P}(t_n^*\geq x)/\Phi_c(x)\geq f(x)\exp(-(1+x)^2/2n(1-2/n)^2)\geq f(x)e^{-2/n(1-2/n)^2}.$$
Note that function $h(x)=x^2/2-x+\ln(2+x)$ takes on its minimum in $[0;1]$ at $x_*=(\sqrt{5}-1)/2\approx 0.618$. Hence $\frac{2}{e}(2+x)\exp(x^2/2-x)>1.256$. Thus,
$$\mathbb{P}(t_n^*\geq x)/\Phi_c(x)>1.25e^{-2/n(1-2/n)^2}\,. \tag{13*}$$
In particular, $\mathbb{P}(t_n^*\geq x)/\Phi_c(x)>1.01$ if $n>12$.

We consider now the case where $x\in[1;\sqrt{n}\,]$. It is well-known that
$$\frac{1}{1+x}<\frac{\Phi_c(x)}{\varphi(x)}<\frac{1}{x}\qquad (x>0), \tag{19}$$
where $\varphi=\Phi'$. Relations (17) – (19) yield
$$\mathbb{P}(t_n^*\geq x)/\Phi_c(x)\geq\mathbb{P}(t_n^*\geq\sqrt{n})/\Phi_c(x)\geq(1-p)^n x/\varphi(x).$$

Let $p=1/n$. Then
$$\mathbb{P}(t_n^*\geq x)/\Phi_c(x)\geq\tfrac{\sqrt{2\pi}}{e}xe^{x^2/2-1/2(n-2)}\,. \tag{20}$$
Since $\inf_{x\geq1}xe^{x^2/2}=e^{1/2}$, we have
$$\mathbb{P}(t_n^*\geq x)/\Phi_c(x)\geq\tfrac{\sqrt{2\pi}}{\sqrt{e}}e^{-1/2(n-2)}\,.$$
Note that $\sqrt{2\pi/e}>1.52$. Thus, (1) and (3) hold. Relation (4) follows from (20). □

Remark 1. The statement of Theorem 1 can be reformulated for negative x by switching from $\{X_i\}$ to $\{-X_i\}$: (3) holds with "$x\geq0$" replaced with "$x\leq0$". Similarly one can reformulate the statement of Theorem 2: as $n\to\infty$,
$$\inf_{x\leq0}\sup_{\mathcal{P}_n}\Big|\mathbb{P}(t_n^*\leq x)/F_n(x)-1\Big|\geq 1/4+o(1). \tag{5*}$$

Remark 2. Distribution (8) is not the only one that can be used in order to establish (1). For instance, let τ, ξ, η be independent r.v.s, $\mathcal{L}(\tau) = \mathbf{B}(c/n)$, where $c \geq 0$, $\mathcal{L}(\xi) = \mathbf{B}(p)$, $\mathbb{E}\eta = 0$, $\mathbb{E}\eta^2 = 1$. Set

$$X = \tau\eta + (1-\tau)(p-\xi)/\sqrt{pq}\,,$$

and let $\{X_i\}$ be independent copies of X. Then $\mathbb{E}X = 0$, $\mathbb{E}X^2 = 1$.

Let, for example, $x = 0$. If $p = 1/n$, then

$$\mathbb{P}(t_n^* \geq 0)/\Phi_c(0) \geq (1-c/n)^n q^{n-1}(q+np) \sim 2/e^{1+c}$$

as $n \to \infty$. Therefore, $\mathbb{P}(t_n^* \geq 0)/\Phi_c(0) \geq 4/e^{1+c} + o(1) > 1$ for all large enough n if $c < \ln(4/e)$.

Proof of Theorem 2 involves Lemma 4 and the argument from the proof of Theorem 1. In view of (6*) it suffices proving the corresponding relations with t_n replaces with t_n^*.

Since $t_n^* \leq \sqrt{n}$, (5) trivially holds if $x_n > \sqrt{n}$. Therefore, we may assume below that $x \in [0; \sqrt{n}\,]$.

Let $X, X_1, ..., X_n$ be defined as in the proof of Theorem 1. Recall that

$$F_n'(x) = C_n(1+x^2/n)^{-(n+1)/2} \qquad (x \in \mathbb{R}),$$

where

$$C_n = \Gamma((n+1)/2)/\sqrt{\pi n}\,\Gamma(n/2),\ \Gamma(y) = \int_0^\infty t^{y-1}e^{-t}dt \qquad (y > 0).$$

We consider first the case where $x \in [1; \sqrt{n}\,]$. Using (22), we derive

$$\begin{aligned} \mathbb{P}(t_n^* \geq x)/F_n^c(x) &\geq \mathbb{P}(t_n^* \geq \sqrt{n})/F_n^c(x) \\ &\geq (1-1/n)^{n+1} x(1+x^2/n)^{(n-1)/2}/C_n \end{aligned} \tag{21}$$

if $p = 1/n$. It is known that $C_n \to 1/\sqrt{2\pi}$ as $n \to \infty$. Since $\inf_{x \geq 1} x(1+x^2/n)^{(n-1)/2} = (1+1/n)^{(n-1)/2} = \sqrt{e} + o(1)$ as $n \to \infty$, (21) yields

$$\mathbb{P}(t_n^* \geq x)/F_n^c(x) \geq \sqrt{2\pi/e} + o(1) \qquad (n \to \infty)$$

uniformly in $x \in [1; \sqrt{n}\,]$.

We consider now the case where $x \in [0; 1]$. Let $p = p_x$. Then

$$\mathbb{P}(t_n^* \geq x) \geq (2+x)e^{-1-x}(1+o(1)) \qquad (n \to \infty)$$

uniformly in $x \in [0; 1]$. Taking into account (2), we notice that $F_n^c(x) - \Phi_c(x) = O(1/n)$ as $n \to \infty$ uniformly in $x \in [0; 1]$. Therefore, (13*) yields

$$\mathbb{P}(t_n^* \geq x)/F_n^c(x) \geq \mathbb{P}(t_n^* \geq x)/\Phi(x)(1+O(1/n)) \geq 1.25 + O(1/n) \qquad (n \to \infty)$$

uniformly in $x \in [0; 1]$. Thus, $\inf_{x \geq 0} \sup_{\mathcal{P}_n} \left| \mathbb{P}(t_n^* \geq x)/F_n^c(x) - 1 \right| \geq 1/4 + o(1)$ as $n \to \infty$.

If $\{x_n\}$ is a non-decreasing sequence of positive numbers such that $1 \ll x_n \leq \sqrt{n}$ as $n \to \infty$, then (21) entails (5*). The proof is complete. □

Lemma 4 *As* $n > 1, x > 0$,

$$\frac{\sqrt{2\pi}\,C_n}{\sqrt{1+1/n}}\Phi_c\Big(x\sqrt{1+1/n}\Big) \leq F_n^c(x) \leq \frac{C_n}{(1-1/n)x}(1+x^2/n)^{-(n-1)/2}\ . \tag{22}$$

Note that (22) means $F_n^c(x)$ decays rather fast when $x\in(0;\sqrt{n}\,]$:

$$F_n^c(x) \le C_n e^{-\frac{x^2}{4}(1-1/n)}/x(1-1/n), \tag{22'}$$

$$F_n^c(x) \ge C_n e^{-\frac{x^2}{2}(1+1/n)}/(1+x)(1+1/n). \tag{22''}$$

Proof of Lemma 4. It is easy to see that

$$\begin{aligned} F_n^c(x) &= C_n \int_x^\infty (1+y^2/n)^{-(n+1)/2}\, dy \\ &\le C_n x^{-1} \int_x^\infty (1+y^2/n)^{-(n+1)/2}\, ydy \\ &= \frac{C_n}{(1-1/n)}\, x^{-1}(1+x^2/n)^{-(n-1)/2} \ . \end{aligned}$$

Using Taylor's formula, one can check that

$$y \ge \ln(1+y) \ge y - y^2/2 \qquad (y\ge 0). \tag{23}$$

Hence

$$e^{x^2} \ge (1+x^2/n)^n \ge \exp(x^2 - x^4/2n) \ge e^{x^2/2} \qquad (0\le x^2\le n). \tag{24}$$

Therefore,

$$F_n^c(x) \le \frac{C_n}{(1-1/n)} x^{-1} e^{-x^2(1-1/n)/4} \,.$$

Similarly,

$$\begin{aligned} F_n^c(x) &\ge C_n \int_x^\infty \exp(-y^2(1+1/n)/2)dy \\ &= C_n \sqrt{2\pi/(1+1/n)}\, \Phi_c\Big(x\sqrt{1+1/n}\,\Big) \\ &\ge C_n e^{-x^2(1+1/n)/2}/(1+x)(1+1/n) \end{aligned}$$

by (19). The proof is complete. □

Proof of Proposition 3. Recall that r.v.s $\{X_i\}$ obey (8^*) and

$$t_n^* = (np - S_n^\xi)/\sqrt{np^2 + (q-p)S_n^\xi}\ ,$$

where $S_n^\xi = \sum_{i=1}^n \xi_i$. Note that

$$t_n^* = g(S_n^\xi),\ Y_n = g(\pi_{np}), \tag{25}$$

where monotone function g is given by (15).

Theorem 4.12 in [5] states that

$$d_{TV}(S_n^\xi;\pi_{np}) \le 3p/4e + 2\delta^2 + 2\delta^*\varepsilon_n, \tag{26}$$

where $\varepsilon_n = \min\Big\{1; (2\pi[(n-1)p])^{-1/2} + 2(1-e^{-np})p/(1-1/n)\Big\}$, $\delta = (1-e^{-np})p,\ \delta^* = (1-e^{-np})p^2$.

Given an arbitrary $A\subset\mathbb{Z}_+$, set $B = g(A)$. Taking into account (25), we observe that

$$\mathbb{P}(t_n^*\in A) - \mathbb{P}(Y_n\in A) = \mathbb{P}(g(S_n^\xi)\in B) - \mathbb{P}(g(\pi_{np})\in B) \le d_{TV}(S_n^\xi;\pi_{np}).$$

Thus, (9) follows from (26). The proof is complete. □

Conclusion. We have shown that the T-test in its present form can be misleading even if the sample size is arbitrarily large: normal or Student's approximation to the distribution of Student's statistics t_n is not automatically applicable if i.i.d.r.v.s $X, X_1, ..., X_n$ are bounded and $\mathrm{var}X=1$.

The paper suggests a generalisation of the T-test that involves checking for the appropriate approximating distribution, and requires estimates of the accuracy of approximation to $\mathcal{L}(t_n)$ with explicit constants. The list of possible approximating distributions may include, beyond normal, functions of Poisson, compound Poisson, and possibly some other infinitely divisible laws.

Acknowledgements

The author is grateful to the reviewer and the Associate Editor for helpful remarks.

References

[1] Arak T.V. and Zaitsev A.Yu. (1986) Uniform limit theorems for sums of independent random variables. — Proc. Steklov Inst. Math., v. 174, 3–214.
[2] Bahadur R.R. (1971) Some limit theorems in statistics. — Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.
[3] Lehmann E.L. (1986) Testing statistical hypotheses. — New York: Springer.
[4] Mason D.M. (2005) The asymptotic distribution of self-normalized triangular arrays. — J. Theoret. Probab., v. 18, No 4, 853–870.
[5] Novak S.Y. (2011) *Extreme value methods with applications to finance.* — London: Chapman & Hall/CRC Press/Taylor & Frensis. ISBN 9781439835746
[6] Novak S.Y. (2018) Non-parametric lower bounds and information functions. — In: Proceedings ISNPS-2016, 69–83. Switzerland: Springer.
[7] Novak S.Y. (2019) Poisson approximation. — *Probability Surveys*, **16**, 228–276; (2021) v. 18, 272–275.
[8] Novak S.Y. (2020) On the T-test. arXiv:submit/3535372
[9] Petrov V.V. (1995) *Limit theorems of Probability Theory.* — Oxford: Clarendon Press.
[10] Pinelis I.F. (2015) Exact bounds on the closeness between the Student and standard normal distributions. — ESAIM Probab. Stat., v. 19, 24–27.
[11] Pinelis I.F. (2016) Optimal-order bounds on the rate of convergence to normality in the multivariate delta method. — arXiv:0906.0177v5
[12] Rukhin A.L. (1993) On the Bahadur efficiency of a t-test. — Sankhyā Indian J. Statistics Ser. A, v. 55, No 1, 159–163.
[13] Shevtsova I.G. (2013) On absolute constants in the Berry–Esseen inequality. — Inform. Appl., v. 7, No 1, 124–125.
[14] Tchebychef P.L. (1886) On integral residue providing approximate values of integrals. — Collective works, v. 2, 444–478.