How normal can the $t$-statistic possibly be?

Thorsten Dickhaus$^1$ and Helmut Finner$^2$

$^1$Department of Mathematics
Humboldt-University Berlin

$^2$Institute of Biometrics & Epidemiology
German Diabetes Center at the Heinrich-Heine-University
Leibniz Center for Diabetes Research, Düsseldorf

Workshop on Validation in Statistics and Machine Learning, WIAS Berlin, 07.10.2010
Outline

Introduction

Chung’s method

Rates of convergence

Expansion in terms of Student’s t

Conclusions
Prologue

Normalized sums are ubiquitous in statistics (binomial models, location parameter tests, \( U \)-statistics, etc.)

Central Limit Theorem (plus Slutsky’s theorem):

\[ (\zeta_n : n \in \mathbb{N}) \text{ iid sequence, } \text{Var}[\zeta_1] < \infty, \bar{\zeta} = \sum_{i=1}^{n} \frac{\zeta_i}{n}, \text{ then} \]

\[ \mathcal{L} \left( \sqrt{n} \frac{\bar{\zeta} - \mathbb{E}[\zeta_1]}{s} \right) \xrightarrow{\text{(n\to\infty)}} \mathcal{N}(0, 1), \quad s = \left( \frac{1}{n-1} \sum_{i=1}^{n} (\zeta_i - \bar{\zeta})^2 \right)^{1/2}. \]

We call \( \sqrt{n}(\bar{\zeta} - \mathbb{E}[\zeta_1])/s \) \( t \)-statistic of \((\zeta_j)_{j=1,...,n}\).

Berry-Esséen:

\( \mathbb{E}[|\zeta_1|^3] \) finite, then rate of convergence is at least \( O(1/\sqrt{n}) \).
Questions in practice

1. Can convergence behavior be characterized more sharply?
2. What roles do higher moments play?
3. Are there means of speeding convergence up?
4. ”How valid” is statistical inference based on the CLT?

Answers (at the end of this talk):

1. YES!
2. A crucial role.
3. YES!
4. It depends.
Edgeworth expansion for standardized sums

Let \((\zeta_n : n \in \mathbb{N})\) iid sequence, \(\mathbb{E}[\zeta_1] = 0\) and \(\text{Var}[\zeta_1] = 1\).

\[
S_n = \sqrt{n} \frac{\zeta}{s} \quad \text{with} \quad \zeta = \sum_{i=1}^{n} \frac{\zeta_i}{n} \quad \text{and} \quad s = \left( \frac{1}{n} \sum_{i=1}^{n} (\zeta_i - \bar{\zeta})^2 \right)^{1/2}.
\]

Modern notation of an (Edgeworth) expansion for the cdf of \(S_n\):

\[
F_n(x) = \Phi(x) + \varphi(x) \sum_{i=1}^{r} n^{-i/2} q_i(x) + o(n^{-r/2}), \quad (1)
\]

with \(\Phi\) cdf and \(\varphi\) pdf of \(\mathcal{N}(0, 1)\). Each \(q_i\) is a polynomial of order \(3i - 1\) with coefficients depending on \(\alpha_j = \mathbb{E}\zeta_1^j, j = 3, \ldots, i + 2\).

Validity (cf. [5]): \(\mathbb{E}|\zeta_1|^{r+2} < \infty\) and Cramér’s condition holds.
The polynomials $q_1$ and $q_2$

The first two polynomials $q_1, q_2$ are computed for example in [5], [6] and can be found in various textbooks. They are given by

$$q_1(y) = \left( \frac{1}{6} + \frac{1}{3}y^2 \right) \alpha_3,$$

$$q_2(y) = \left( \frac{1}{12}y^3 - \frac{1}{4}y \right) \alpha_4 + \left( \frac{1}{6}y - \frac{1}{18}y^5 - \frac{1}{9}y^3 \right) \alpha_3^2 - \frac{1}{2}y^3.$$

This representation shows that the rate of convergence is $O(n^{-1/2})$ in case of $\alpha_3 \neq 0$ and $O(n^{-1})$ in case of $\alpha_3 = 0$.

Obviously, the best possible rate of convergence is $O(n^{-1})$ and this may be the reason that usually only the first two polynomials are reported.
Impact of normalization

Lehmann and Romano (2005), Section 11.4.2:
Edgeworth expansion for the classical $t$-statistic with normalization $(n - 1)^{-1}$ in the definition of $s$.

Approximation polynomials in this case differ from the $q_i$’s in (1).

Hence, the norming sequence in the denominator of a self-normalized sum is of importance for its asymptotic behavior.

⇒ Question: Exist other norming sequences for specific values of the moments $\alpha_i, i \geq 3$, such that the rate of convergence can be improved?

Back in 1946, Kai-Lai Chung derived an expansion for $F_n$.

Unfortunately, the explicit expansion given in equation (35) in [2] is incorrect as noted earlier by Wallace in [9] and to our knowledge there seems to be no published correction.

We corrected the main inaccuracy in [2] and extended the formulas where necessary.

⇒ Chung’s method elementary, straightforward and efficient!

In principle, the $q_i$’s are computable up to arbitrary order with an algebraic computer package.
Correction of Chung’s error

In the derivations in [2], the function $g$ defined by

$$g(\lambda) = z(1 + \lambda^2 z^2)^{-1/2} \left[ 1 + \sum_{j=1}^{\infty} \frac{\Gamma(3/2)}{\Gamma(3/2 - j)\Gamma(j + 1)} (\alpha_4 - 1)^{j/2} (\lambda x)^j \right]$$

and its derivatives $g^{(i)}$ play a crucial role.

The formulas given in [2], p. 458, struggle by abbreviating $z' = z(1 + \lambda^2 z^2)^{-1/2}$ and ignoring that $z'$ depends on $\lambda$.

Correct derivatives in $\lambda = 0$ are given by

$$g^{(1)}(0) = \frac{1}{2} z (\alpha_4 - 1)^{1/2} x, \quad g^{(2)}(0) = -z^3 - \frac{1}{4} z (\alpha_4 - 1) x^2,$$

$$g^{(3)}(0) = -\frac{3}{2} z^3 (\alpha_4 - 1)^{1/2} x + \frac{3}{8} z (\alpha_4 - 1)^{3/2} x^3,$$

$$g^{(4)}(0) = 9 z^5 + \frac{3}{2} z^3 (\alpha_4 - 1) x^2 - \frac{15}{16} z (\alpha_4 - 1)^2 x^4.$$
Chung’s approximation technique

Formally, $F_n(z)$ is approximated in [2] by

$$
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\lambda) \ w(x, y) \ dy \ dx + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \gamma(x, y) \ dy \ dx,
$$

with $w(x, y) = \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}\right)$, $\rho = \alpha_3(\alpha_4 - 1)^{-1/2}$.

For the definition of $\gamma(x, y)$, we need some more notation:

$$
w_{pq}(x, y) = \frac{\partial^{p+q}}{\partial x^p \partial y^q} w(x, y), \quad I_{pq}(z) = \int_{-\infty}^{\infty} x^r w_{pq}(x, z) \ dx,
$$

$$
f_{pq}(\lambda) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\lambda) \ w_{pq}(x, y) \ dy \ dx, \quad h(\zeta) = t_1 \frac{\zeta^2 - 1}{(\alpha_4 - 1)^{1/2}} + t_2 \zeta,
$$

where $\zeta$ has the same distribution as $\zeta_1$. 
Derivation of $\gamma(x, y)$

Let $U_j(t_1, t_2)$ denote the $j$th cumulant of $h(\zeta)$ and define

$$m_k(t_1, t_2) = \sum_{\ell=0}^{k-3} \frac{-i^{\ell+1} U_{\ell+3}(t_1, t_2)}{\ell + 3)!} \chi^{\ell+1},$$

$$\Psi_k(it_1, it_2) = \sum_{j=1}^{k-3} \frac{m_j(t_1, t_2)^j}{j!}.$$

Expanding the $U_i$’s in terms of $t_1, t_2$ and replacing $(it_1)^p(it_2)^q$ by $(-1)^{p+q}w_{pq}(x, y)$ in $\Psi_k(it_1, it_2)$ yields the representation

$$\Psi_k(it_1, it_2) \equiv \sum_{j=1}^{k-3} \gamma_j(x, y) = \gamma(x, y),$$

where $\gamma_j(x, y) = O(\lambda^j)$ and $w_{pq}(x, y)$ appears repeatedly in $\gamma_j(x, y)$ for various $p, q$ with $p + q \leq 3r.$
Taylor expansion for $f_{pq}(\lambda)$

We can write

$$F_n(z) = \sum_{j=0}^{r} \frac{f_{00}^{(j)}(0)}{j!} + \sum_{j=1}^{r} \int_{-\infty}^{\infty} \int_{-\infty}^{g(\lambda)} \gamma_j(x, y) \, dy \, dx + o(\lambda^r).$$

Now, $f_{pq}(\lambda)$ is approximated by the Taylor series in $\lambda = 0$ wherever it appears in $\int_{-\infty}^{\infty} \int_{-\infty}^{g(\lambda)} \gamma_j(x, y) \, dy \, dx$. More precisely, $f_{pq}(\lambda)$ is replaced by

$$\sum_{j=0}^{r} \frac{f_{pq}^{(j)}(0)}{j!} \lambda^j + o(\lambda^r).$$

This means, all we need to carry out Chung’s method are tractable formulas for $f_{pq}^{(j)}(0)$!
Lemma:
For $q \geq 0$, non-vanishing $I_{pq}$’s are given by the following recursion.

\[
\begin{align*}
I_{0q}^0(z) &= \varphi^{(q)}(z), \\
I_{0q}^1(z) &= -\rho \varphi^{(q+1)}(z), \\
I_{0q}^{r+1}(z) &= -\rho I_{0,q+1}^r(z) + rI_{0q}^{r-1}(z), r \geq 1, \\
I_{pq}^r(z) &= -rI_{p-1,q}^{r-1}(z), 1 \leq p \leq r.
\end{align*}
\]

Remark:
Modified Hermite polynomials: $h_n(x) = -(I/\sqrt{2})^nH_n(Ix/\sqrt{2})$.
Interestingly, $I_{0q}^r = h_r(\rho D)(\varphi^{(q)})$, where $D$ denotes the differential operator. Note that (2) corresponds to $h_{r+1}(x) = -x h_r(x) + r h_{r-1}(x)$.

$(X, Z)$ bivariate normal, means 0, variances 1 and covariance $\rho$:

\[
I_{0q}^r(z) = \frac{\partial^q}{\partial z^q} \mathbb{E}[X^r | Z = z].
\]
Lemma: (Computation of $f_{pq}^{(j)}(0)$)

Setting $I_{pq}^r \equiv I_{pq}^r(z)$ for $r = 0, \ldots, 3$, we have for $p, q \geq 0$

\[
\begin{align*}
  f_{00}(0) & = \Phi(z), \\
  f_{pq}(0) & = \begin{cases} 
- I_{p,q-1}^0, & q \geq 1, \\
0, & q = 0,
\end{cases} \\
  f_{pq}^{(1)}(0) & = \frac{1}{2} z (\alpha_4 - 1)^{1/2} I_{pq}^1, \\
  f_{pq}^{(2)}(0) & = \frac{1}{4} (\alpha_4 - 1) (-z I_{pq}^2 + z^2 I_{pq+1}^2) - z^3 I_{pq}^0, \\
  f_{pq}^{(3)}(0) & = -\frac{3}{2} (\alpha_4 - 1)^{1/2} (z^3 I_{pq}^1 + z^4 I_{p,q+1}^1) \\
& \quad + \frac{1}{8} (\alpha_4 - 1)^{3/2} (3z I_{pq}^3 - 3z^2 I_{p,q+1}^3 + z^3 I_{p,q+2}^3).
\end{align*}
\]
Computational remarks

Now all ingredients for the computation of the polynomials $q_i$ are collected.

Computation by hand remains cumbersome. Therefore we prepared a Maple worksheet which allows the computation of the $q_i$’s up to arbitrary order.

Due to the structure of the $f_{pq}^{(j)}$’s, the lemma can be extended by utilizing standard symbolic integration methods.

Clearly, resources limit the number of computable $q_i$’s.

Remark:
We also computed the $q_i$’s with Hall’s ’smooth function’ method described in [6] up to order 6 with complete coincidence. Hall’s method involves the computation of moments of more complicated statistics and seems more time consuming.
Rates of convergence

Recall that $q_1(y) \equiv 0$ for $\alpha_3 = 0$.

**Interpretation:** Vanishing skewness of $\zeta_1 \Rightarrow$ On the $n^{-1/2}$ scale, the approximation of $F_n$ cannot be distinguished from $\Phi$.

However, the rate of convergence of $F_n$ towards $\Phi$ can at most be $O(n^{-1})$, because $q_2$ never vanishes.

**Our approach to improve this rate of convergence:**

$$T_n = \frac{\sqrt{n} \bar{\zeta}}{\sqrt{a_n \sum_{i=1}^{n} (\zeta_i - \bar{\zeta})^2}}, \quad \text{where} \quad a_n = \frac{1}{n(1 - \sum_{j=1}^{M} C_j n^{-j/2})}$$

Formal expansion for the generalized self-normalized sum $T_n$:

$$F_{T_n}(t) = \Phi(t) + \sum_{i=1}^{r} n^{-i/2} \tilde{q}_i(t) \varphi(t) + o(n^{-r/2}) \quad (3)$$

Coefficients of the $\tilde{q}_i$’s depend on cumulants of $\zeta_1$ and on $C_j$’s.
Derivation of the approximation for $T_n$

Notice that $T_n = S_n/b_n$ with $b_n = \sqrt{na_n}$. Therefore,

$$F_{T_n}(t) = \mathbb{P}(T_n \leq t) = \mathbb{P}(S_n \leq b_nt) = F_n(b_nt) .$$

From (1), we get under the necessary moment condition that

$$F_{T_n}(t) = \Phi(b_nt) + \sum_{i=1}^{r} n^{-i/2} q_i(b_nt) \varphi(b_nt) + o(n^{-r/2}) .$$

In terms of $\Phi(t)$ and $\varphi(t)$, we can write

$$F_{T_n}(t) = \Phi(t) + \varphi(t) \left[ h_n(t) + \sum_{i=1}^{r} n^{-i/2} q_i(b_nt) g_n(t) \right] ,$$

where the auxiliary functions $h_n$ and $g_n$ are defined by

$$h_n(t) = \left[ \frac{\Phi(b_nt)}{\Phi(t)} - 1 \right] \frac{\Phi(t)}{\varphi(t)} , \quad g_n(t) = \frac{\varphi(b_nt)}{\varphi(t)} .$$

$\Rightarrow$ Expansions for $b_n$, $h_n(t)$, $g_n(t)$ needed!
Lemma:

Setting $\lambda = n^{-1/2}$, asymptotic expansions of $b_n$, $h_n(t)$ and $g_n(t)$ are given by

\[
b_n = 1 + \frac{C_1}{2} \lambda + \frac{C_2 + 3C_1^2/4}{2} \lambda^2 + O(\lambda^3),
\]

\[
h_n(t) = \frac{C_1 t}{2} \lambda + \frac{t}{8} \left(4C_2 + 3C_1^2 - C_1^2t^2\right) \lambda^2 + O(\lambda^3),
\]

\[
g_n(t) = 1 - \frac{C_1 t^2}{2} \lambda - \frac{t^2}{2} \left(C_1^2 + C_2 - \frac{C_1^2 t^2}{4}\right) \lambda^2 + O(\lambda^3).
\]

Proof:

The expansions for $b_n$ and $g_n(t)$ are simple applications of the Taylor series of the square root and the exponential functions. For the expansion of $h_n(t)$, well-known asymptotic expansions for Mills’ ratio are needed additionally.
Resulting approximation polynomials

Having expanded $b_n$, $h_n(t)$, and $g_n(t)$ in this manner, we finally obtain the first two $\tilde{q}_i$’s as

$$\tilde{q}_1(t) = \frac{\alpha_3 t^2}{3} + \frac{\alpha_3}{6} + \frac{C_1 t}{2},$$

$$\tilde{q}_2(t) = \frac{3t C_1^2}{8} + \frac{\alpha_4 t^3}{12} + \frac{\alpha_3^2 t}{6} - \frac{t^3 C_1^2}{8} - \frac{\alpha_3^2 t^3}{9} - \frac{\alpha_3^2 t^5}{18}$$
$$+ \frac{\alpha_3 C_1 t^2}{4} + \frac{t C_2}{2} - \frac{t^3}{2} - \frac{\alpha_3 C_1 t^4}{6} - \frac{\alpha_4 t}{4}.$$
Sanity check

Setting $M = 2$, $C_1 = 0$ and $C_2 = 1$, we get the Studentized sum

$$
\tilde{S}_n = \sqrt{n} \tilde{\zeta} / \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (\zeta_i - \bar{\zeta})^2}
$$

with corresponding approximation polynomials

$$
\tilde{q}_1(t) = \frac{\alpha_3}{3} t^2 + \frac{\alpha_3}{6} = \frac{\alpha_3}{6} (2t^2 + 1),
$$

$$
\tilde{q}_2(t) = \frac{\alpha_4 t^3}{12} + \frac{\alpha_3^2 t}{6} - \frac{\alpha_3^2 t^3}{9} - \frac{\alpha_3^2 t^5}{18} + \frac{t}{2} - \frac{t^3}{2} - \frac{\alpha_4 t}{4}
$$

$$
= t \left[ \frac{\kappa}{12} (t^2 - 3) - \frac{\alpha_3^2}{18} (t^4 + 2t^3 - 3) - \frac{1}{4} (t^2 + 1) \right],
$$

where $\kappa = \alpha_4 - 3$ denotes the excess kurtosis of $\zeta_1$. These are just the approximation polynomials given in Section 11.4.2 of the textbook [8] by Lehmann and Romano.
Rates of convergence for generalized self-normalized sums

**Theorem:**
Let \( \Delta_n(x) = |F_{T_n}(x) - \Phi(x)|. \)

(i) If \( \alpha_3 \neq 0 \) or \( C_1 \neq 0 \), then \( \Delta_n(x) = O(n^{-1/2}) \).

(ii) If \( \alpha_3 = C_1 = 0 \) and \( (\alpha_4 \neq 6 \) or \( C_2 \neq 3) \), then \( \Delta_n(x) = O(n^{-1}) \).

(iii) If \( \alpha_3 = C_1 = 0 \) and \( \alpha_4 = 6 \) and \( C_2 = 3 \) and \( (\alpha_5 \neq 0 \) or \( C_3 \neq 0) \), then \( \Delta_n(x) = O(n^{-3/2}) \).

(iv) If \( \alpha_3 = C_1 = 0 \) and \( \alpha_4 = 6 \) and \( C_2 = 3 \) and \( \alpha_5 = C_3 = 0 \), then \( \Delta_n(x) = O(n^{-2}) \).

**Proof:**
For parts (i)-(iii), we subsequently solve \( \tilde{q}_i(t) \equiv 0 \) for \( C_i \) and \( \alpha_{i+2} \) for \( i = 1, 2, 3 \). For the proof of part (iv), we show that it is impossible to find values for \( (\alpha_6, C_4) \) such that \( \tilde{q}_4 \) becomes the null polynomial.
Remark:

1. The studentized sum $\tilde{S}_n$ with $C_i = 0$ for all $i \neq 2$ and $C_2 = 1$ can only achieve a rate of convergence of $O(n^{-1})$.

2. Justification for the special role of $C_2 = 3$:
   Norming $a_n = (n - 3)^{-1}$ leads to variance standardization of $T_n$, that is, $\text{Var}[T_n] = 1$ if the $\zeta_i$ are iid normal as $\mathcal{N}(0, 1)$.

3. The special role of $\alpha_4 = 6$ in parts (ii) - (iv) is not clear to us.

Example:

Let $\varphi(x|\sigma)$ denote the pdf of $\mathcal{N}(0, \sigma^2)$ and the density of $\zeta_1$ given by

$$f_{\zeta_1}(x) = \alpha \varphi(x|\sigma_1) + (1 - \alpha) \varphi(x|\sigma_2)$$

with $\sigma_1^2 = (2\alpha)^{-1}$, $\sigma_2^2 = (2(1 - \alpha))^{-1}$ and $\alpha = (2 + \sqrt{2})/4$.

Then $\mathbb{E}\zeta_1 = \mathbb{E}\zeta_1^3 = \mathbb{E}\zeta_1^5 = 0$, $\mathbb{E}\zeta_1^2 = 1$, $\mathbb{E}\zeta_1^4 = 6$, and $\mathbb{E}\zeta_1^6 = 90$.

Setting $C_1 = C_3 = 0$ and $C_2 = 3$, part (iv) applies in this case.
Expansion in terms of Student’s $t$

Goal here: Derive an Edgeworth-type expansion for $T_n$ of form

$$F_{T_n}(t) = F_{t_{\nu}}(t) + \sum_{i=1}^{r} n^{-i/2} Q_i(t) \varphi(t) + o(n^{-r/2}) \quad (4)$$

in terms of Student’s $t$ with $\nu = n - 1$ degrees of freedom.

Note: $T_n$ with norming sequence $a_n = (n - 1)^{-1}$ and $\zeta_1 \sim \mathcal{N}(0, 1)$ is exactly $t_{\nu}$-distributed.

Questions:

1. Can an improved rate of convergence be obtained by changing the approximating distribution from $\mathcal{N}(0, 1)$ to $t_{\nu}$?

2. Can the norming constants $C_j$ be employed to correct for higher-order moments of $\zeta_1$?
Derivation of the $Q_i$’s in (4)

Denote by $q_i^*, i = 1, \ldots, r$, the approximation polynomials for $T_n = t_\nu$, i.e., for special choices $M = 2$, $C_1 = 0$ and $C_2 = 1$ and $\alpha_j, j = 3, \ldots, (r + 2)$, equal to the moments of $\mathcal{N}(0, 1)$.

By subtracting the resulting expansion from the general expansion for $T_n$, we immediately conclude that

$$Q_i(t) = \tilde{q}_i(t) - q_i^*(t), i = 1, \ldots, r.$$ 

Carrying out these calculations, we obtain the first four $q_i^*$’s as

$$q_1^*(t) = q_3^*(t) \equiv 0,$$

$$q_2^*(t) = -\frac{t}{4} \left( t^2 + 1 \right),$$

$$q_4^*(t) = -\frac{t}{96} \left( 3 t^6 - 7 t^4 + 19 t^2 + 21 \right).$$
Consequently, the first two $Q_i$'s are given by

$$Q_1(t) = \tilde{q}_1(t) = \frac{\alpha_3}{3} t^2 + \frac{C_1}{2} t + \frac{\alpha_3}{6},$$

$$Q_2(t) = -\frac{\alpha_3^2}{18} t^5 - \frac{\alpha_3 C_1}{6} t^4 - \left(\frac{1}{4} + \frac{\alpha_3^2}{9} - \frac{\alpha_4}{12} + \frac{C_1^2}{8}\right) t^3 + \frac{\alpha_3 C_1}{4} t^2$$

$$+ \left(\frac{3 C_1^2}{8} + \frac{C_2}{2} + \frac{\alpha_3^2}{6} - \frac{\alpha_4}{4} + \frac{1}{4}\right) t.$$

Rates of convergence:

- $Q_1$ only vanishes for $\alpha_3 = C_1 = 0$.
- $Q_2$ only vanishes if additionally $\alpha_4 = 3$ and $C_2 = 1$, i.e., in case of coincidence with the classical $t$-distribution case.
- This need for coincidence extends to the conditions for vanishing $Q_3$ and $Q_4$ (explicit formulas omitted here).
- **Conclusion:** $t$-approximation instead of normal approximation does not help to increase convergence rates.
A link to Gayen’s (1949) method

Substitute $\varphi$ in (4) by the pdf $f_{t_{\nu}}$ of the $t_{\nu}$-distribution:

$$F_{Tn}(t) = F_{t_{\nu}}(t) + \sum_{i=1}^{r} n^{-i/2} \tilde{Q}_i(t) f_{t_{\nu}}(t) + o(n^{-r/2}).$$

(5)

Closely related expressions for $F_{Tn}$ for fixed $n$ have already been investigated in 1949 by A. K. Gayen based on M. S. Bartlett’s famous paper [1].

One can derive the first four $\tilde{Q}_i$’s in (5) by expanding

$$\varphi(t) = f_{t_{\nu}}(t) \left[ 1 + \frac{1 + 2t^2 - t^4}{4n} + O(n^{-2}) \right].$$

Plugging the latter expansion into (4) leads to

$$\tilde{Q}_i \equiv Q_i, \ i = 1, 2,$$

$$\tilde{Q}_i(t) = Q_i(t) + \frac{1 + 2t^2 - t^4}{4} Q_{i-2}(t), \ i = 3, 4.$$
Comparison with Gayen’s results

Unfortunately, we could only reproduce Gayen’s (1949) results up to order $n^{-1}$.

Taking limits ($n \to \infty$) in Gayen’s paper also yields $\tilde{Q}_i \equiv Q_i$, $i = 1, 2$.

However, the expressions of order $O(n^{-3/2})$ associated with the factors $\alpha_3^2$ and $\alpha_3 \alpha_4$ seem to be in error in [4], p. 359, and also taking limits ($n \to \infty$) in these expressions does not coincide with our results.

Therefore, we also recomputed the original approximation method by Bartlett (cf. [1]) which underlies Gayen’s (1949) calculations and finally reproduced “our” $\tilde{Q}_i$’s for $i = 1, \ldots, 4$. 
Asymptotic order of magnitude of $|F_{T_n} - F_{t_{\nu}}|$

Utilizing Chung’s method and higher order expansions for $\varphi(t)/f_{t_{\nu}}(t)$, we calculated $Q_i$’s and $\tilde{Q}_i$’s up to order 8.

$$\alpha_k^* : k - \text{th moment of } \mathcal{N}(0, 1), \Delta \alpha_k = \alpha_k^* - \alpha_k$$

$$C_2^* = 1, C_k^* = 0 \text{ for } k \neq 2 \text{ and } \Delta C_k = C_k^* - C_k$$

**Corollary:**

Assume that the $(M + 2)$-nd moment $\alpha_{M+2}$ of $\zeta_1$ is finite for some integer $1 \leq M \leq 8$ and Cramér’s condition holds. Then

$$|F_{T_n} - F_{t_{\nu}}| = O(n^{-k^*/2}),$$

where $k^* = \min\{k \in \{1, \ldots, M\} : \Delta \alpha_{k+2} \neq 0 \lor \Delta C_k \neq 0\}$. If no such $k^*$ exists, then $|F_{T_n} - F_{t_{\nu}}| = o(n^{-M/2})$. 
What happens for $M > 8$?

Since each polynomial $Q_i$ or $\tilde{Q}_i$, respectively, only depends on $\alpha_j$, $j = 3, \ldots, i + 2$ and $C_j, j = 1, \ldots, i$, and equations (4), (5) are valid for $T_n = t_\nu$, it is clear that also for $M > 8$ the conditions

$$\Delta \alpha_{i+2} = 0 \land \Delta C_i = 0 \text{ for all } i = 1, \ldots, M$$  \hspace{1cm} (6)

imply $Q_i(t) \equiv 0$ and $\tilde{Q}_i(t) \equiv 0$ for all $i = 1, \ldots, M$.

$\Rightarrow$ Conditions (6) are sufficient for vanishing polynomials up to the $M$-th for arbitrary $M \in \mathbb{N}$.

We conjecture that conditions (6) are also necessary conditions for any $M \geq 1$ as stated in the Corollary for $1 \leq M \leq 8$. 
Conclusions

• Four different types of Edgeworth expansions for $S_n$, $T_n$. Once polynomials for one are obtained, they can be utilized to derive the polynomials for the others straightforwardly.

• At http://www.helmut-finner.de, find Maple sheets for Chung’s, Hall’s, and the Bartlett-Gayen method.

• Practical implications:
  • For skewed distributions, no convergence rate improvement upon $O(n^{-1/2})$ is possible with our approach.
  • If there is any evidence that $\alpha_4$ is near 6, a normal approximation with $C_2 = 3$ is the best choice leading to $|F_{T_n}(x) - \Phi(x)| = O(n^{-3/2})$.
  • $t$-approximation works best for $a_n = (n - 1)^{-1}$ and can achieve arbitrary rate of convergence for universes which are close to standard normal in terms of moments. This makes the $t$-approximation a more natural choice if we assume that the universe is ”nearly normal”.


