1

t-Tests With Models Close To The Normal Distribution

Alfonso García-Pérez

Departamento de Estadística, Investigación Operativa y Cálculo Numérico
Facultad de Ciencias
Universidad Nacional de Educación a Distancia
Paseo Senda del Rey 9, 28040-Madrid (Spain)
e-mail: agar-per@ccia.uned.es

Abstract: The t-distribution is a very usual distribution for several test statistics because a normal distribution is frequently assumed as underlying model. Even in some tests based on robust statistics, such as the test based on the sample trimmed mean, a t-distribution is used as distribution for the standardized sample trimmed mean if the underlying model is normal. Nevertheless, it is necessary a deeper understanding of the behaviour of these kind of tests and the computations of their key elements, such as the p-value and the critical value, with small samples, when the underlying model is close but different from the normal distribution. In this paper we obtain good analytic approximations with small samples, of the p-value and the critical value of a t-test (i.e., a test with a t-distribution for the test statistic under a normal model), studying its behaviour when the underlying distribution is close but different from the normal model. We conclude the paper studying some robustness properties of t-tests.

Keywords and phrases: Robustness in hypotheses testing, von Mises expansion, tail area influence function, saddlepoint approximation, robustness of t-tests

1.1 Introduction

Many classical parametric tests were obtained assuming a normal distribution as underlying model. This is the reason why the χ^2, Student's t-, and F-distributions play a prominent role in Statistics as distributions for test statistics.

Also in some tests based on robust statistics, such as the test based on the α-trimmed mean, a t-distribution is used as distribution for the standardized trimmed mean if the underlying model is normal, even with a small sample size; see, for instance, Tukey and McLaughlin (1963), Staudte and Sheather (1990) or Wilcox (1997). See also Patel et al. (1988).

Nevertheless, it is necessary a deeper understanding of the behaviour of these kind of tests and the computations of their key elements, such as the p-value and the critical value, with small samples, when the underlying distribution is not normal but a slight deviation from it. Previous studies are, for instance, Benjamini (1983), Cressie (1980), Chen and Loh (1990) or Sawilowsky and Blair (1992). Really, the distribution of the Student's statistic under other models is usually obtained through simulations, except in the paper by Lee and Gurland (1977) where this distribution is obtained only under contaminated normal models.

Here we obtain good closed form approximations of some key elements of a t-test, such as the critical value and the p-value, in a close to normal situation, developing a method proposed in García-Pérez (2003), which is based on considering all these elements as functionals of the model distribution, and that makes use of the von Mises expansion of a functional plus, in some cases, saddlepoint approximations.

This method is specially useful in robustness studies where the model distribution is, frequently, a slight deviation from the normal distribution (for instance, a contaminated normal) but complicated enough to render impossible an exact calculation of these elements.

With these aims, in Section 2 we briefly explain the method that we will use in the following sections.

We obtain, in Section 3, von Mises approximations for t-tests, (i.e., tests in which the test statistic follows a t-distribution under a normal model), but now when the model distribution is close to the normal distribution.

In Section 4 we obtain saddlepoint approximations of the von Mises approximations and some interesting results; for instance, a complementary result of the obtained by Benjamini (1983) or the conclusions drawn by Cressie (1980), which is that "a light-tailed parent distribution causes a heavy-tailed t-distribution". We also study the robustness of t-tests, obtaining some results that confirm the idea that, also with small samples sizes, the t-test has robustness of validity, at least in the tails, with slight departures from the normality.

1.2 Preliminaries

Although the method that we are going to explain in this paper can be extended to a more general setting, we will consider in it a one-dimensional test based on a test statistic $T_n = T_n(X_1, ..., X_n)$ that rejects the null hypothesis H_0 when

T_n is larger than the critical value k_n^F and where F is the distribution that the X_i's follow under H_0. If $T_n = t$, the p-value will be then the tail probability $p_n^F = P_F\{T_n > t\}$.

In particular, we will consider t-tests in the paper, i.e., tests in which T_n follows a t-distribution under a normal model, studying here its behaviour under a model F, close but different, from the normal.

In these tests we will just consider two elements, the critical value k_n^F and the p-value p_n^F, although the method can be used to approximate other elements like the power. One of the key points is to consider these elements as functionals of the model distribution F.

We will suppose that T_n is real valued although the sample $X_1, ..., X_n$ can be one- or multi-dimensional. The only restriction is that, under the null hypothesis, both the critical value k_n^F and the p-value p_n^F must be functionals of only one distribution function F that we will assume univariate.

In a one-dimensional parametric test of the null hypothesis $H_0 : \theta = \theta_0$, if $X_1, ..., X_n$ is a sample from a random variable X with distribution function F_θ and $F_{n;\theta}$ is the cumulative distribution function of the test statistic T_n, the critical value of the level-α test

$$k_n^F = F_{n;\theta_0}^{-1}(1-\alpha)$$

and the p-value

$$p_n^F = P_{F_{\theta_0}}\{T_n > t\}$$

will be considered functionals of F_{θ_0}. (Throughout the paper, the inverse of any distribution function G is defined, as usual, by $G^{-1}(s) = \inf\{y | G(y) \geq s\}$, $0 < s < 1$.)

For instance, if $T_n = M$ is the sample median, then

$$k_n^F = F_{\theta_0}^{-1}(B^{-1}(1-\alpha))$$

and

$$p_n^F = 1 - B(F_{\theta_0}(t))$$

where B is the cumulative distribution function of a beta $\beta((n+1)/2 , (n+1)/2)$.

If $T_n = \overline{x}$ is the sample mean and $F_{\theta_0} \equiv \Phi_{\theta_0,\sigma}$ the normal distribution $N(\theta_0, \sigma)$, it is

$$k_n^F = \frac{1}{\sqrt{n}} \left(\Phi_{\theta_0,\sigma}^{-1}(1-\alpha) + \theta_0 \left(\sqrt{n} - 1\right)\right)$$

and the p-value

$$p_n^F = 1 - \Phi_{\theta_0,\sigma}\left(t\sqrt{n} - \theta_0(\sqrt{n}-1)\right).$$

The most common t-test is the based on the usual t-statistic

$$T_n = \frac{\sqrt{n}(\overline{x} - \theta_0)}{S}$$

with a t_{n-1} distribution under a $N(\theta_0, \sigma)$ model.

Besides, if $k = [n\alpha]$ is the integer portion of $n\alpha$, another t-test is the standardized sample α-trimmed mean (removing the k largest and k smallest observations)

$$T_n = \frac{(1-2\alpha)\sqrt{n}(\overline{x}_\alpha - \mu_{\alpha,0})}{S_w}$$

with an approximate t_{n-2k-1} distribution under a normal model, where S_w^2 is the sample Winsorized variance, if the null hypothesis is about the parameter α-trimmed mean, $H_0 : \mu_\alpha = \mu_{\alpha,0}$. See Tukey and McLaughlin (1963), Wilcox (1997, p. 75), or Staudte and Sheather (1990, p. 105, 156, 186). See also the paper by Patel et al. (1988).

Finally, let us observe that it does not matter that the functionals k_n^F and p_n^F depend on n because we are not interested in the asymptotic (in n) distribution properties of these functionals. Actually, n is in both of them what Reeds (1976, p.39) calls *an auxiliary parameter*.

1.2.1 Influence functions of p_n^F and k_n^F

To obtain the von Mises expansions of the functionals p_n^F and k_n^F we will need their influence functions with respect to a model G (that later we will assume it to be the normal distribution).

We will represent these influence functions, respectively, as $\dot{p}_n^G$ and $\dot{k}_n^G$; they will be based on the *Tail Area Influence Function* (TAIF) defined by Field and Ronchetti (1985). This one is just the influence function of the tail probability of a statistic T_n at a distribution G and it is defined as

$$\text{TAIF}(x; t; T_n, G) = \frac{\partial}{\partial\epsilon} P_{G^\epsilon}\{T_n > t\}\bigg|_{\epsilon=0}$$

for all $x \in I\!R$ where the right hand side exists, being $G^\epsilon := (1-\epsilon)G + \epsilon\delta_x$ the contaminated model, and δ_x the point mass distribution at $x \in I\!R$.

The TAIF is really the influence function of the p-value,

$$\dot{p}_n^G = \text{TAIF}(x; t; T_n, G)$$

and after some computations (see García-Pérez, 2003, for details) it is

$$\overset{\bullet}{k}{}_n^G = \frac{\mathrm{TAIF}(x; k_n^G; T_n, G)}{g_n(k_n^G)}$$

assuming that the distribution function G_n of the test statistic, under the model G, has a density g_n with respect to the Lebesgue measure and that $g_n(k_n^G) \neq 0$.

Since, in this paper, the distribution G (called *pivotal distribution* in the sequel) will be the normal distribution, we will have no problem about this with $\overset{\bullet}{k}{}_n^G$.

1.2.2 Von Mises expansions of p_n^F and k_n^F

Let T be a functional defined on a convex set $\mathcal{F}$ of distribution functions and with range the set of the real numbers.

If F and G are two members of $\mathcal{F}$ and $s \in [0,1]$ is a real number, let us define the function A of the real variable s by

$$A(s) = T\left((1-s)G + sF\right) = T\left(G + s(F-G)\right).$$

Considering the viewpoint adopted by Filippova (1961) and Reeds(1976), the (*low-brow way* of the) von Mises expansion of the functional T is just the ordinary Taylor expansion of the real function $A(s)$, assuming that A satisfies the usual conditions for a Taylor expansion to be valid if $s \in [0,1]$; see, for instance, Serfling (1980, p. 43, theorem 1.12.1A).

Then, expanding $A(s)$ about $s = 0$ and evaluating the resultant expansion at $s = 1$, we obtain the *von Mises expansion of the functional* T *at the distribution* $F \in \mathcal{F}$

$$T(F) = T(G) + \sum_{k=1}^{m} \frac{A^{k)}(0)}{k!} + Rem \qquad (1.1)$$

where $A^{k)}(0)$ is the ordinary kth derivative of A at the point 0,

$$A^{k)}(0) = \left.\frac{d^k}{dt^k} A(t)\right|_{t=0} \qquad k = 1, ..., m$$

and where the remainder term Rem depends on F and G, and on the $(m+1)$th derivative of A (i.e., on the influence function of T, if there exists).

Considering the sum in (1.1) up to the first or second term, we have, respectively, the *first-order von Mises expansion*

$$T(F) = T(G) + A^{1)}(0) + Rem_1$$

and the *second-order von Mises expansion*

$$T(F) = T(G) + A^{1)}(0) + \frac{1}{2}\, A^{2)}(0) + Rem_2$$

of T, having the second one a higher degree of accuracy than the first one. Because we will obtain very accurate approximations just considering the first-order expansion, we will always consider this one in the rest of the paper, omitting in the sequel the subscript of the remainder term.

Moreover, if there exists the influence function of the functional T, usually represented by $\overset{\bullet}{T}(x)$ or just by $IF(x;T,G)$, it is

$$T(F) = T(G) + \int IF(x;T,G)\, dF(x) + Rem$$

being the remainder term

$$Rem = \frac{1}{2} \int \int T_H(x,y)\, dF(x) dF(y)$$

where

$$T_H(x,y) = \left. \frac{\partial}{\partial \epsilon} IF(x;T,H_{\epsilon,y}) \right|_{\epsilon=0} + IF(y;T,H)$$

and $H(x) = G(x) + \lambda(F(x) - G(x))$ with λ some constant in $[0,1]$ depending on F, G, T, and $H_{\epsilon,y} = (1-\epsilon)H + \epsilon\, \delta_y$ the H-contaminated distribution. (See García-Pérez, 2003, for more details.)

Then, if there exists $\overset{\bullet}{p}{}_n^G$, the (first-order) *von Mises expansion of the p-value* will be

$$p_n^F = p_n^G + \int \overset{\bullet}{p}{}_n^G(x)\, dF(x) + Rem$$

and, if there exists $\overset{\bullet}{k}{}_n^G$, the (first-order) *von Mises expansion of the critical value* will be

$$k_n^F = k_n^G + \int \overset{\bullet}{k}{}_n^G(x)\, dF(x) + Rem$$

where the remainder terms, usually different in both expansions, will be smaller as F and G are closer. This can be formalized with the usual sup-norm or with a tail ordering on distributions like the $<_t$-ordering defined by Loh (1984).

1.2.3 Von Mises approximations of p_n^F and k_n^F with a model F close to the normal distribution

From the previous von Mises expansions we define the approximations we were looking for, using the normal distribution $\Phi_{\mu,\sigma}$, as distribution G.

So, we define the (first-order) *von Mises (VOM) approximation of* p_n^F *by* p_n^Φ as

$$p_n^F \simeq p_n^\Phi + \int \mathrm{TAIF}(x; t; T_n, \Phi_{\mu,\sigma})\, dF(x) \tag{1.2}$$

and the (first-order) *von Mises (VOM) approximation of* k_n^F *by* k_n^Φ as

$$k_n^F \simeq k_n^\Phi + \frac{1}{\phi_n(k_n^\Phi)} \int \mathrm{TAIF}(x; k_n^\Phi; T_n, \Phi_{\mu,\sigma})\, dF(x) \tag{1.3}$$

where ϕ_n is the density of T_n under the normal model $\Phi_{\mu,\sigma}$.

In these equations we see explicitly the extra term that we add to the usual asymptotic normal approximations p_n^Φ and k_n^Φ, that improve them.

To simplify the notation we will usually omit the parameters of the normal distribution when it appears as subscript or superscript. We will represent the distribution and density functions of the standard normal $N(0,1)$, respectively, by Φ_s and ϕ_s.

1.3 Von Mises Approximations For t-Tests

In this section we will consider t-tests, i.e., tests such that the test statistic T_n follows a t-distribution under the null hypothesis, when the underlying model is the normal distribution $N(\mu,\sigma)$. Here we will determine the VOM approximations (1.2) and (1.3), for their p-value and critical value, when the underlying model distribution F is not normal but a slight deviation from it.

The key element in the VOM approximations (1.2) and (1.3) is the TAIF under the normal model. To obtain this, we express first the tail probability of a t-test as a functional of the cumulative distribution function $\Phi_{\mu,\sigma}$ of the normal distribution $N(\mu,\sigma)$.

If the test statistic T_n follows a t-distribution with n degrees of freedom, t_n, we can express the tail probability of T_n as

$$P_\Phi\{T_n > t\} = \frac{1}{2} \int_{-\infty}^{\infty} P_\Phi \left\{ \chi_n^2 \leq \frac{n(y-\mu)^2}{t^2\sigma^2} \right\} \, d\Phi_{\mu,\sigma}(y)$$

where χ_n^2 is a random variable with a χ_n^2 distribution.

Therefore, under the contaminated model $\Phi^\epsilon = (1-\epsilon)\,\Phi_{\mu,\sigma} + \epsilon\,\delta_x$, we have

$$P_{\Phi^\epsilon}\{T_n > t\} \quad = \quad \frac{1}{2} \left((1-\epsilon) \int_{-\infty}^{\infty} \left[1 - P_{\Phi^\epsilon} \left\{ \chi_n^2 > \frac{n(y-\mu)^2}{t^2\sigma^2} \right\} \right] \, d\Phi_{\mu,\sigma}(y) \right.$$

$$+\epsilon\left[1-P_{\Phi^\epsilon}\left\{\chi^2_n>\frac{n(x-\mu)^2}{t^2\sigma^2}\right\}\right]\Bigg).$$

Now, we express the $\mathrm{TAIF}(x;t;t_n,\Phi_{\mu,\sigma})$ in terms of the TAIF of the χ^2, $\mathrm{TAIF}(x;t;\chi^2_n,\Phi_{\mu,\sigma})$ as

$$\begin{aligned}\mathrm{TAIF}(x;t;t_n,\Phi_{\mu,\sigma}) &= \frac{\partial}{\partial\epsilon}P_{\Phi^\epsilon}\{T_n>t\}\Big|_{\epsilon=0}\\ &= \frac{1}{2}\left(-\int_{-\infty}^{\infty}P\left\{\chi^2_n\le\frac{n(y-\mu)^2}{t^2\sigma^2}\right\}\,d\Phi_{\mu,\sigma}(y)\right.\\ &\quad -\int_{-\infty}^{\infty}\mathrm{TAIF}\left(x;\frac{n(y-\mu)^2}{t^2\sigma^2};\chi^2_n,\Phi_{\mu,\sigma}\right)d\Phi_{\mu,\sigma}(y)\\ &\quad \left.+P\left\{\chi^2_n\le\frac{n(x-\mu)^2}{t^2\sigma^2}\right\}\right).\end{aligned}$$

In García-Pérez (2004) we obtained that the TAIF, under a normal model, of the functional χ^2_n test considered is, if $n>1$,

$$\mathrm{TAIF}(x;t;\chi^2_n,\Phi_{\mu,\sigma})=n\,P\left\{\chi^2_{n-1}>t-\left(\frac{x-\mu}{\sigma}\right)^2\right\}-n\,P\{\chi^2_n>t\}.$$

Then, if $n>1$, the TAIF under a normal model, of the functional t-test considered is

$$\begin{aligned}\mathrm{TAIF}(x;t;t_n,\Phi_{\mu,\sigma}) &= \frac{n}{2}-(n+1)P\{t_n>t\}+\frac{1}{2}P\left\{\chi^2_n\le\frac{n(x-\mu)^2}{t^2\sigma^2}\right\}\\ &\quad -\frac{n}{2}\int_{-\infty}^{\infty}P\left\{\chi^2_{n-1}>\frac{n(y-\mu)^2}{t^2\sigma^2}-\frac{(x-\mu)^2}{\sigma^2}\right\}\,d\Phi_{\mu,\sigma}(y).\end{aligned}$$

Because in García-Pérez (2004) we also obtained that it is

$$\int_{-\infty}^{\infty}P\left\{\chi^2_{n-1}>t-\frac{(x-\mu)^2}{\sigma^2}\right\}\,d\Phi_{\mu,\sigma}(x)=P\{\chi^2_n>t\}$$

it is easy to check out that it is

$$\int_{-\infty}^{\infty}\mathrm{TAIF}(x;t;t_n,\Phi_{\mu,\sigma})\,d\Phi_{\mu,\sigma}(x)=0.$$

Finally, to obtain the VOM approximation of the p-value p_n^F and the critical value k_n^F when the model distribution for the observable random variable is F, we have to integrate the last TAIF with respect to F, as it is showed in expressions (1.2) and (1.3), obtaining that

$$\begin{aligned} p_n^F \;\simeq\;& \frac{n}{2} - n\,P\{t_n > t\} + \frac{1}{2}\int_{-\infty}^{\infty} P\left\{\chi_n^2 \leq \frac{n(x-\mu)^2}{t^2\sigma^2}\right\} dF(x) \\ & -\frac{n}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P\left\{\chi_{n-1}^2 > \frac{n(y-\mu)^2}{t^2\sigma^2} - \frac{(x-\mu)^2}{\sigma^2}\right\} d\Phi_{\mu,\sigma}(y)\, dF(x) \end{aligned} \tag{1.4}$$

and

$$\begin{aligned} k_n^F \;\simeq\;& t_{n;\alpha} + \frac{1}{g_{t_n}(t_{n;\alpha})}\left[\frac{n}{2} - (n+1)\,\alpha + \frac{1}{2}\int_{-\infty}^{\infty} P\left\{\chi_n^2 \leq \frac{n(x-\mu)^2}{t_{n;\alpha}^2\,\sigma^2}\right\} dF(x)\right. \\ & \left. -\frac{n}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P\left\{\chi_{n-1}^2 > \frac{n(y-\mu)^2}{t_{n;\alpha}^2\,\sigma^2} - \frac{(x-\mu)^2}{\sigma^2}\right\} d\Phi_{\mu,\sigma}(y)\, dF(x)\right] \end{aligned} \tag{1.5}$$

where $t_{n;\alpha}$ is the $(1-\alpha)$-quantile of a t_n distribution and g_{t_n} the density function of this distribution.

Example 1.3.1 (t-tests under a scale contaminated normal model) *If we suppose a sample from a scale contaminated normal model $F = (1-\epsilon)N(\mu,\sigma) + \epsilon N(\mu, k\,\sigma)$, the VOM p-value (1.4) and VOM critical value (1.5) of a t_n test are, respectively,*

$$\begin{aligned} p_n^F \;\simeq\;& P\{t_n > t\} + \epsilon\left[\frac{n}{2} - (1+n)\,P\{t_n > t\} + P\{t_n > t/k\}\right. \\ & \left. -\frac{n}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P\left\{\chi_{n-1}^2 > \frac{n(y-\mu)^2}{t^2\sigma^2} - \frac{(x-\mu)^2}{\sigma^2}\right\} d\Phi_{\mu,\sigma}(y)\, d\Phi_{\mu,k\sigma}(x)\right] \end{aligned}$$

and

$$\begin{aligned} k_n^F \;\simeq\;& t_{n;\alpha} + \frac{\epsilon}{g_{t_n}(t_{n;\alpha})}\left[\frac{n}{2} - (1+n)\,\alpha + P\{t_n > t_{n;\alpha}/k\}\right. \\ & \left. -\frac{n}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P\left\{\chi_{n-1}^2 > \frac{n(y-\mu)^2}{t_{n;\alpha}^2\sigma^2} - \frac{(x-\mu)^2}{\sigma^2}\right\} d\Phi_{\mu,\sigma}(y)\, d\Phi_{\mu,k\sigma}(x)\right] \end{aligned}$$

Table 1.1: *Exact* and approximate p-values under a contaminated normal model and $n = 4$

t	"exact"	VOM
1	$0'1979$	$0'1979$
2	$0'0688$	$0'0716$
4	$0'0138$	$0'0158$
5	$0'0071$	$0'0089$

We see in these two expressions the extra term we add to the usual p-value and critical value of a t-test under a normal model, and the influence of each element ($\epsilon, n, k, ...$) *in these extra terms.*

To finish the example with numerical values, let us consider, for instance, a sample of size 4 from a distribution $0'95\, N(0,1) + 0'05\, N(0, \sqrt{4})$ *instead of a* $N(0,1)$*, and the usual Student's test statistic*

$$\frac{\sqrt{4}\,\overline{x}}{S}$$

that follows a t_3 *distribution, to test at level* α*,* $H_0 : \mu = 0$ *against* $H_1 : \mu > 0$*. The approximated p-value and critical value are, respectively,*

$$\begin{aligned} p_n^F &\simeq P\{t_3 > t\} + 0'05 \left[\frac{3}{2} - 4\, P\{t_3 > t\} + P\{t_3 > t/2\}\right. \\ &\left. -\frac{3}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P\left\{\chi_2^2 > \frac{3\, y^2}{t^2} - x^2\right\} d\Phi_{0,1}(y)\, d\Phi_{0,2}(x)\right] \end{aligned}$$

and

$$\begin{aligned} k_n^F &\simeq t_{3;\alpha} + \frac{0'05}{g_{t_3}(t_{3;\alpha})} \left[\frac{3}{2} - 4\, \alpha + P\{t_3 > t_{3;\alpha}/2\}\right. \\ &\left. -\frac{3}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P\left\{\chi_2^2 > \frac{3\, y^2}{t_{3;\alpha}^2} - x^2\right\} d\Phi_{0,1}(y)\, d\Phi_{0,2}(x)\right] \end{aligned}$$

In Tables 1.1 and 1.2, appear the exact values *(obtained through simulation of a 30000 samples and using the package 'stepfun' of the software R in different t's) and the (first-order) VOM approximations in this situation (using the package 'adapt' of R for the numerical integration).*

Table 1.2: *Exact* and approximate critical values under a contaminated normal model and $n = 4$

α	"exact"	VOM
$0'01$	$4'445$	$4'718$
$0'05$	$2'330$	$2'378$
$0'1$	$1'629$	$1'631$

Nevertheless, to obtain these results we had to compute the approximations using numerical integration. Usually we would prefer to have analytic expressions for them that can be used as elements in other more complex problems; for instance, to study the robustness of the t-tests. For this reason we will obtain saddlepoint approximations of these (first-order) VOM approximations in the next section.

1.4 Saddlepoint approximations for t-tests

Although it is possible to use known saddlepoint approximations for the tails of the χ^2 and t-distributions that appear in expressions (1.4) and (1.5) —see, for instance, Jensen (1995, p. 49, 86)—, these ones would be numerical again, or would depend on integrals of the normal cumulative distribution function with respect to the underlying model F, not obtaining, in this way, manageable analytic expressions of them. For this reason we will approximate the TAIF, using the Lugannani and Rice formula, before integration in (1.2) and (1.3).

If T_n follows a t_n distribution, and Y_1, Y_2 are two independent gamma distributions, respectively $\gamma(1/2, 1/2)$ and $\gamma(n/2, n/2)$, we can write

$$P\{T_n > t\} = P\{Y_1 - t^2 Y_2 > 0\}$$

where the random variable $Y = Y_1 - t^2 Y_2$ has cumulant generating function

$$K(\theta) = \log M(\theta) = \log M_\gamma(\theta) + n \log M_\gamma(-\theta t^2/n)$$

being

$$M_\gamma(\theta) = \int_{-\infty}^{\infty} e^{\theta\,(u-\mu)^2/\sigma^2}\, d\Phi_{\mu,\sigma}(u)$$

the moment generating function of a gamma $\gamma(1/2, 1/2)$, a functional that depends on the distribution model $\Phi_{\mu,\sigma}$.

Now, we can use the Lugannani and Rice formula —see Lugannani and Rice (1980) or, better, Daniels (1983)—, for the tail, in a sample of size one, of $Y = Y_1 - t^2 Y_2$, obtaining that it is

$$P_\Phi\{Y>0\} = 1-\Phi_s(w)+\phi_s(w)\left\{\frac{1}{r}-\frac{1}{w}+O(1)\right\} \tag{1.6}$$

where the functionals r and w are

$$w = sign(z_0)\sqrt{-2\,K(z_0)}$$

$$r = z_0\sqrt{K''(z_0)}$$

that depend on the saddlepoint z_0, which is the solution of the equation

$$K'(z_0)=0$$

from where we obtain the saddlepoint

$$z_0 = \frac{t^2-1}{2\,t^2}\,\frac{n}{1+n}.$$

Now, from (1.6), we obtain

$$\begin{aligned}
\text{TAIF}\,(x;t;t_n,\Phi_{\mu,\sigma}) &= \frac{\partial}{\partial\epsilon}P_{\Phi^\epsilon}\,\{Y>0\}\Big|_{\epsilon=0} \\
&\simeq \frac{\phi_s(w)}{r}\left\{-w\;\dot{w}-\frac{\dot{r}}{r}+\frac{\dot{w}\;r}{w^2}\right\} \\
&= \frac{e^K}{\sqrt{2\pi}\,z_0\sqrt{K''}}\left\{\dot{K}\left[1-\frac{z_0\sqrt{K''}}{(-2K)^{3/2}}\right]-\frac{\dot{z}_0}{z_0}-\frac{\dot{K}''}{2K''}\right\}
\end{aligned}$$

After some computations and approximations, and if it is

$$A_1 = \frac{t}{\sqrt{\pi}\,(t^2-1)}\,e^{-(t^2-1)/2}$$

and

$$A_2 = 1-\frac{t^2-1}{\sqrt{2}(t^2-1-2\log t)^{3/2}}$$

we obtain $\forall x$ and $t>1$ for the functional t-test considered, that it is

$$\text{TAIF}\,(x;t;t_n,\Phi_{\mu,\sigma}) = A_1\left\{\left(A_2-\frac{3t^2+1}{4(t^2-1)}\right)t^{-1}\,e^{(t^2-1)(x-\mu)^2/(2t^2\sigma^2)}\right.$$

$$+\frac{3t^2-1}{2(t^2-1)}\, t^{-3} \left(\frac{x-\mu}{\sigma}\right)^2 e^{(t^2-1)(x-\mu)^2/(2t^2\sigma^2)}$$
$$-\frac{t^{-5}}{4}\left(\frac{x-\mu}{\sigma}\right)^4 e^{(t^2-1)(x-\mu)^2/(2t^2\sigma^2)}$$
$$\left.-\frac{t^2}{t^2-1}\left(\frac{x-\mu}{\sigma}\right)^2+\frac{t^2}{t^2-1}-A_2\right\}.$$

(It is easy to check out that it is $\int_{-\infty}^{\infty} \text{TAIF}\,(x;t;t_n,\Phi_{\mu,\sigma})\;d\Phi_{\mu,\sigma}(x)=0$).

Now, integrating this TAIF with respect to a model F, from (1.2) we obtain the *VOM+SAD approximate p-value* of the test, under a model F,

$$p_n^F \simeq P\{t_n>t\}+\int \text{TAIF}(x;t;t_n,\Phi_{\mu,\sigma})\,dF(x). \tag{1.7}$$

From (1.3), we obtain that the *VOM+SAD approximate critical value* of the test, under a model F, is

$$k_n^F \simeq t_{n;\alpha}+\frac{1}{g_{t_n}(t_{n;\alpha})}\int \text{TAIF}(x;t_{n;\alpha};t_n,\Phi_{\mu,\sigma})\,dF(x) \tag{1.8}$$

where $t_{n;\alpha}$ is the $(1-\alpha)$-quantile of a t_n distribution with density g_{t_n}.

Example 1.4.1 (t-tests under scale contaminated normal models) *Let us consider a t-test, i.e., a test in which the test statistic follows a t_n distribution under a normal model $N(0,1)$. Now, let us consider as model for this test, a scale contaminated normal model $F=(1-\epsilon)N(0,1)+\epsilon N(0,k)\equiv(1-\epsilon)\,\Phi_s+\epsilon\,\Phi_{0,k}$.*

Because, if $t>1$ and $a=0,2,$ or 4, it is

$$\int_{-\infty}^{\infty} x^a\, e^{(t^2-1)x^2/(2t^2)}\, d\Phi_{0,k}(x)=\frac{2^{a/2}\,\Gamma((a+1)/2)\,t^{a+1}\,k^a}{\sqrt{\pi}\,[t^2-k^2(t^2-1)]^{(a+1)/2}}$$

if $k<\sqrt{1/(1-t^{-2})}$, the VOM+SAD p-value (1.7) is

$$\begin{aligned} p_n^F \;\simeq\;& P\{t_n>t\}+\epsilon\,A_1\left\{\left(A_2-\frac{3t^2+1}{4(t^2-1)}\right)[t^2-k^2(t^2-1)]^{-1/2}\right.\\ &+\frac{3t^2-1}{t^2-1}\,\frac{k^2}{2}\,[t^2-k^2(t^2-1)]^{-3/2}-\frac{3\,k^4}{4}\,[t^2-k^2(t^2-1)]^{-5/2}\\ &\left.+\frac{t^2(1-k^2)}{t^2-1}\;-A_2\right\}\end{aligned}$$

Table 1.3: *Exact* and approximate p-values under a scale contaminated normal model, and standard p-values. $n = 10$

t	"exact"	VOM+SAD	$P\{t_9 > t\}$
$1'5$	$0'08533$	$0'09908$	$0'08393$
2	$0'03787$	$0'03978$	$0'03828$
3	$0'00760$	$0'00749$	$0'00748$

and the VOM+SAD critical value (1.8),

$$\begin{aligned} k_n^F \;\simeq\; & t_{n;\alpha} + \frac{\epsilon\, A_1}{g_{t_n}(t_{n;\alpha})} \left\{ \left(A_2 - \frac{3t_{n;\alpha}^2 + 1}{4(t_{n;\alpha}^2 - 1)} \right) [t_{n;\alpha}^2 - k^2(t_{n;\alpha}^2 - 1)]^{-1/2} \right. \\ & + \frac{3t_{n;\alpha}^2 - 1}{t_{n;\alpha}^2 - 1} \frac{k^2}{2} [t_{n;\alpha}^2 - k^2(t_{n;\alpha}^2 - 1)]^{-3/2} - \frac{3\,k^4}{4} [t_{n;\alpha}^2 - k^2(t_{n;\alpha}^2 - 1)]^{-5/2} \\ & \left. + \frac{t_{n;\alpha}^2(1 - k^2)}{t_{n;\alpha}^2 - 1} - A_2 \right\} \end{aligned}$$

where, as before, $t_{n;\alpha}$ is the $(1-\alpha)$-quantile of a t_n distribution with density g_{t_n}.

Now, if we consider a scale contaminated normal model $0'95\,N(0,1) + 0'05\,N(0,0'6)$ (a situation with inliers), a sample of size $n = 10$ and the usual Student's test statistic

$$\frac{\sqrt{n}\,\overline{x}}{S}$$

that follows a t_9 distribution under a normal model, to test at level α, $H_0 : \mu = 0$ against $H_1 : \mu > 0$, the VOM+SAD p-values and critical values are shown in Tables 1.3 and 1.4 together with the exact ones (obtained with a simulation of 30000 samples and using the package 'stepfun' of R), and the usual values obtained under a standard normal model $N(0,1)$.

From these tables we observe that the VOM+SAD approximations are quite good and, comparing these ones with the last column (values under a normal model) we see that, for most of the values, using a light-tailed model we obtain a long-tailed distribution for the test statistic.

Remark 1.4.1 One of the questions related with the behaviour of t-tests is if they are conservative or liberal with long-tailed and short-tailed distributions, i.e. that, if it is $F >_t G$, with $>_t$ a (partial) ordering of distribution functions then, is it $P_G\{t_n > t\} \geq P_F\{t_n > t\}$?

Table 1.4: *Exact* and approximate critical values under a scale contaminated normal model, and standard critical values. $n = 10$

α	"exact"	VOM+SAD	$t_{9;\alpha}$
$0'01$	$2'82284$	$2'82323$	$2'82144$
$0'05$	$1'84097$	$1'87263$	$1'83311$
$0'1$	$1'39572$	$1'58127$	$1'38303$

Table 1.5: Actual sizes of the one-sample t-test when sampling from two scale contaminated normal models. $n = 3$

Nominal level of significance (α)	$0'98N(0,1) + 0'02N(0,0'6)$	$0'95N(0,1) + 0'05N(0,0'6)$
$0'01$	$0'00999$	$0'00999$
$0'05$	$0'05013$	$0'05031$
$0'1$	$0'10306$	$0'10764$

A complete answer to this question depends on the integrals of the TAIF with respect to the distribution model through expression (1.7). Thanks to the previous example we obtain a solution inside the class of scale contaminated normal models, complementary of the conclusion drawn by Benjamini (1983), which is that "a light-tailed parent distribution causes a heavy-tailed t-distribution". Namely,

If we consider two distributions $F_{k_1} = (1-\epsilon)\,\Phi_s + \epsilon\,\Phi_{0,k_1}$ and $F_{k_2} = (1-\epsilon)\,\Phi_s + \epsilon\,\Phi_{0,k_2}$ where $0 < k_1 < k_2 < 1$, (i.e., $F_{k_2} >_t F_{k_1}$ with respect, for instance, to the tail ordering defined by Loh, 1984) it is

$$P_{F_{k_1}}\{t_n > t\} \geq P_{F_{k_2}}\{t_n > t\}$$

at least if the critical value t is $1 < t \leq 1.747$.

Remark 1.4.2 From Table 1.3 we can see that the size of the test does not change very much with a distribution $0'95\,N(0,1) + 0'05\,N(0,0'6)$ considering a t_9-distribution. In Tables 1.5 and 1.6 we obtain the same conclusions with a t_3- and a t_5-distributions, respectively.

From these computations we can state that, with small samples, the t-test has robustness of validity for small departures from a Gaussian population, at least in the tails. Nevertheless, this is not, probably, the most common situation we meet in the real life. Sawilowsky and Blair (1992) agree with both conclusions.

Table 1.6: Actual sizes of the one-sample t-test when sampling from two scale contaminated normal models. $n = 5$

Nominal level of significance (α)	$0'98N(0,1) + 0'02N(0,0'6)$	$0'95N(0,1) + 0'05N(0,0'6)$
$0'01$	$0'00999$	$0'00999$
$0'05$	$0'05056$	$0'05141$
$0'1$	$0'10689$	$0'11723$

Table 1.7: Actual sizes of the one-sample t-test when sampling from location contaminated normal models. $n = 3$

Nominal level of significance (α)	$0'98N(0,1)+$ $0'02N(\mp 0'5,1)$	$0'95N(0,1)+$ $0'05N(\mp 0'5,1)$	$0'9N(0,1)+$ $0'1N(\mp 0'5,1)$
$0'01$	$0'00999$	$0'00997$	$0'00995$
$0'05$	$0'04978$	$0'04945$	$0'04890$
$0'1$	$0'09862$	$0'09655$	$0'09310$

Example 1.4.2 (t-tests under location contaminated normal models) *If we suppose a sample from a location contaminated normal model $F = (1-\epsilon)N(0,1) + \epsilon N(\mu,1)$, the VOM+SAD p-value of a t_n test is*

$$p_n^F \simeq P\{t_n > t\} + \epsilon\, A_1 \left\{ \left(e^{\mu^2(t^2-1)/2} - 1 \right) \left(A_2 + \frac{t^2\mu^2}{t^2-1} \right) - e^{\mu^2(t^2-1)/2}\, \frac{\mu^4 t^4}{4} \right\}.$$

(There is no problem to compute the critical value or to consider others more general location contaminated normal models, or even location-scale contaminated normal models).

Remark 1.4.3 From Table 1.7 we obtain, for location contaminated normal models, the same conclusions than before in the sense that, with small samples, the t-test has robustness of validity, at least in the tails, for small departures from a Gaussian population.

References

1. Benjamini, Y. (1983). Is the t test really conservative when the parent distribution is long-tailed?, *Journal of the American Statistical Association*, **78**, 645–654.

2. Chen, H. and Loh, W. Y. (1990). Uniform robustness against nonnormality of the t and F tests, *Communications in Statistics*, **19**, 3707–3723.

3. Cressie, N. (1980). Relaxing assumptions in the one sample t-test, *Australian Journal of Statistics*, **22**, 143–153.

4. Daniels, H. E. (1983). Saddlepoint approximations for estimating equations, *Biometrika*, **70**, 89–96.

5. Field, C. A. and Ronchetti, E. (1985). A tail area influence function and its application to testing, *Communications in Statistics*, **4**, 19–41.

6. Filippova, A. A. (1961). Mises'theorem on the asymptotic behaviour of functionals of empirical distribution functions and its statistical applications, *Theory of Probability and its Applications*, **7**, 24–57.

7. García-Pérez, A. (2003). Von Mises approximation of the critical value of a test, *Test*, **12**, 385–411.

8. García-Pérez, A. (2004). Chi-square tests under models close to the normal distribution, Technical Report, Departamento de Estadística, Investigación Operativa y Cálculo Numérico, UNED, Madrid.

9. Jensen, J. L. (1995). *Saddlepoint Approximations*, Oxford: Clarendon Press, New York.

10. Lee, A. F. S. and Gurland, J. (1977). One-sample t-test when sampling from a mixture of normal distributions, *The Annals of Statistics*, **5**, 803–807.

11. Loh, W. Y. (1984). Bounds on AREs for restricted classes of distributions defined via tail-orderings, *The Annals of Statistics*, **12**, 685–701.

12. Lugannani, R. and Rice, S. (1980). Saddle point approximation for the distribution of the sum of independent random variables, *Advances in Applied Probability*, **12**, 475–490.

13. Patel, K. R., Mudholkar, G. S. and Fernando, J. L. I. (1988). Student's t approximations for three simple robust estimators, *Journal of the American Statistical Association*, **83**, 1203–1210.

14. Reeds, J. A. (1976). *On the Definitions of Von Mises Functionals*, Ph.D. thesis, Department of Statistics, Harvard University, Cambridge, Massachusetts.

15. Sawilowsky, S. S. and Blair, R. C. (1992). A more realistic look at the robustness and type II error properties of the t test to departures from population normality, *Psychological Bulletin*, **111**, 352–360.

16. Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

17. Staudte, R. G. and Sheather, S. J. (1990). *Robust Estimation and Testing*, Wiley, New York.

18. Tukey, J. W. and McLaughlin, D. H. (1963). Less vulnerable confidence and significance procedures for location based on a single sample: Trimming/Winsorization 1, *Sankhya, A*, **25**, 331–352.

19. Wilcox, R. R. (1997). *Introduction to Robust Estimation and Hypothesis Testing*, Academic Press, San Diego.