symmetry

MDPI

Article

Symmetry, Asymmetry and Studentized Statistics

Maria de Fátima Brilhante [1,2], Dinis Pestana [2,3,4,*] and Maria Luísa Rocha [2,5,6]

1 Faculdade de Ciências e Tecnologia, Universidade dos Açores, Rua da Mãe de Deus, 9500-321 Ponta Delgada, Portugal; maria.fa.brilhante@uac.pt
2 Centro de Estatística e Aplicações, Universidade de Lisboa (CEAUL), Campo Grande, 1749-016 Lisboa, Portugal; maria.ls.rocha@uac.pt
3 Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
4 Instituto de Investigação Científica Bento da Rocha Cabral, Calçada Bento da Rocha Cabral 14, 1250-012 Lisboa, Portugal
5 Faculdade de Economia e Gestão, Universidade dos Açores, Rua da Mãe de Deus, 9500-321 Ponta Delgada, Portugal
6 Centro de Estudos de Economia Aplicada do Atlântico (CEEAplA), Universidade dos Açores, Rua da Mãe de Deus, 9500-321 Ponta Delgada, Portugal
* Correspondence: ddpestana@ciencias.ulisboa.pt

Abstract: Inferences on the location parameter λ in location-scale families can be carried out using Studentized statistics, i.e., considering estimators $\widetilde{\lambda}$ of λ and $\widetilde{\delta}$ of the nuisance scale parameter δ, in a statistic $T = g(\widetilde{\lambda}, \widetilde{\delta})$ with a sampling distribution that does not depend on (λ, δ). If both estimators are independent, then T is an externally Studentized statistic; otherwise, it is an internally Studentized statistic. For the Gaussian and for the exponential location-scale families, there are externally Studentized statistics with sampling distributions that are easy to obtain: in the Gaussian case, Student's classic t statistic, since the sample mean $\widetilde{\lambda} = \overline{X}$ and the sample standard deviation $\widetilde{\delta} = S$ are independent; in the exponential case, the sample minimum $\widetilde{\lambda} = X_{1:n}$ and the sample range $\widetilde{\delta} = X_{n:n} - X_{1:n}$, where the latter is a dispersion estimator, which are independent due to the independence of spacings. However, obtaining the exact distribution of Student's statistic in non-Gaussian populations is hard, but the consequences of assuming symmetry for the parent distribution to obtain approximations allow us to determine if Student's statistic is conservative or liberal. Moreover, examples of external and internal Studentizations in the asymmetric exponential population are given, and an ANalysis Of Spacings (ANOSp) similar to an ANOVA in Gaussian populations is also presented.

Keywords: studentized statistics; symmetric distributions; exponential family; location parameter; scale parameter

check for updates

Citation: Brilhante, M.d.F.; Pestana, D.; Rocha, M.L. Symmetry, Asymmetry and Studentized Statistics. *Symmetry* **2024**, *16*, 1297. https://doi.org/10.3390/sym16101297

Academic Editor: Qing-Wen Wang

Received: 12 September 2024
Revised: 25 September 2024
Accepted: 29 September 2024
Published: 2 October 2024

1. Introduction

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random sample from a population $X \sim \text{Gaussian}(\mu, \sigma)$. If the location parameter μ and the scale parameter σ are unknown, the unbiased estimators $\widetilde{\mu} = \overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k \sim \text{Gaussian}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ and $\widetilde{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2$, with $\frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2_{n-1}$, can be used to estimate these parameters.

The distribution of the estimator $\overline{X}$ depends on the nuisance parameter σ, which is a problem for making inferences on the location parameter μ. Student's [1] pathbreaking paper has shown that the statistic

$$t_{n-1} = \sqrt{n}\,\frac{\overline{X} - \mu}{S} \tag{1}$$

has a distribution that does not depend on the nuisance scale parameter, since its probability density function is

$$f_{t_{n-1}}(t) = \frac{1}{B\left(\frac{1}{2}, \frac{n-1}{2}\right)} \frac{1}{\sqrt{n-1}\left(1+\frac{t^2}{n-1}\right)^{n/2}} \mathbb{I}_{\mathbb{R}}(t)\,, \tag{2}$$

where $B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\mathrm{d}x$, $p,q>0$, is Euler's Beta function. The probability density function defined in (2) is for a random variable with Student's distribution with $n-1$ degrees of freedom. Basically, the "Studentization" defined in (1), as opposed to the common standardization $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$, uses the estimators of the parameter of interest (location) and of the nuisance parameter, so that the sampling distribution of $t_{n-1} = g(\overline{X}, S)$ does not depend on the location parameter μ or on the nuisance parameter σ, thus being a pivot statistic that can be used for inferences on μ.

However, Student's exceptional result depends heavily on the fact that in the location-scale Gaussian family, $\overline{X}$ and S^2 are independent random variables, which is an exclusive property of the Gaussian family. This characterization of the Gaussian distribution appeared in Geary's work [2], and has been proved, independently, by Darmois [3] and by Skitovich [4]. Additionally, in the context of the Koopman–Darmois–Pitman k-parameter exponential family [5–7], the Gaussian family is the only one with support $\mathbb{R}$ and with a pair of sufficient statistics for (μ,σ), namely $\left(\sum_{k=1}^n X_k, \sum_{k=1}^n X_k^2\right)$, which captures all available information in the sample to estimate these two parameters.

When working with samples from non-Gaussian location-scale families, the dependence structure between the sample mean $\overline{X}_n = \frac{1}{n}\sum_{k=1}^n X_k$ and the sum of squares $SS_n = \sum_{k=1}^n (X_k - \overline{X}_n)^2 = (n-1)S_n^2$ is hard to discern, except for the special case $n=2$ (note that an index n is now being used to emphasize the size of the random sample). For $n>2$, the computation of the probability density function of Student's statistic

$$T_{n-1} = \sqrt{n(n-1)}\,\frac{\overline{X}_n - \lambda}{\sqrt{SS_n}} \tag{3}$$

is a more difficult problem to handle. In Section 2, the joint probability density function of $(\overline{X}_n, SS_n)$ is investigated in a general context, i.e., not restricted to the Gaussian family. Its explicit expression for $n=2$ is given in Subsection 2.1, and examples for the probability density function of T_1, i.e., for $n=2$ in (3), for several symmetric parent distributions are given in Subsection 2.2. For $n \geq 3$, a recurrence formula for the joint probability density function of $(\overline{X}_n, SS_n)$ is obtained in Subsection 2.3, and a detailed study of the recurrence formula for the Gaussian case is seen in Subsection 2.4.

An interesting alternative for Studentization was proposed in Logan et al.'s work [8], where a self-normalized statistic $W_{n-1}^* = \frac{\sum_{k=1}^n X_k}{\left(\sum_{k=1}^n |X_k|^\alpha\right)^{1/\alpha}}$ was considered, especially when X has heavier tails than the Gaussian law, and it is in the domain of attraction of a stable law for sums with index $\alpha \in (0,2)$. Peña et al.'s [9] monograph on self-normalized processes is a thorough overview, with Chapter 15 dealing with the classical t-statistic and Studentized statistics.

Another interesting fact is that Efron [10] showed that $T_{n-1} \stackrel{d}{=} t_{n-1}$ when a rotational symmetry of the unit vector $\boldsymbol{U} = \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} = \left(\frac{X_1}{\|\boldsymbol{X}\|}, \ldots, \frac{X_n}{\|\boldsymbol{X}\|}\right)$ over the surface of the unit sphere in the Euclidean n-space E^n is assumed. However, Efron [10] pointed out the following:

> Unfortunately the usual sampling procedures almost never yield rotational symmetry for the normalized vector $\boldsymbol{U}$ except in the case $X \sim \text{Gaussian}(0,\sigma^2)$.

Efron also stated that "A very special "lucky" case is given in Section 9, namely Inverted normal error for $n=2$. It is possible to construct examples where T_n is t-distributed with n degrees of freedom, without $\boldsymbol{U}$ having rotational symmetry". Efron's [10] pioneering work also investigated the consequences, in what regards Student's statistic, by assuming weaker orthant symmetry, i.e., when the random vector $\boldsymbol{U} = (U_1, \ldots, U_n)$ has the same distribution as $\boldsymbol{U}_\delta = (\delta_1 U_1, \ldots, \delta_n U_n)$ for every choice of $\delta_i = \pm 1$, $i = 1, 2, \ldots, n$.

This led us to observe, in Subsection 2.4, that the joint probability density function of $(\overline{X}_n, SS_n)$, in the exceptional rotational symmetry of the Gaussian case, is proportional to the product of n Gaussian$(0, \sigma)$ probability density functions, computed at some special points that form a symmetric arithmetic progression with a null sum and a sum of squares equal to one.

In Section 3, we shall investigate the simplifications that result from an additional symmetry assumption, namely when the joint probability density function of $(\overline{X}_n, SS_n)$ is proportional to the product of n probability density functions of X, as in the Gaussian case. The approximation obtained from the smoothness hypothesis 1 (Subsection 3.2) is further investigated in Subsection 3.3 for the worst case possible, i.e., uniform distribution (see Hendriks et al. [11]), by comparing the approximate expression for T_2 with the exact distribution given by Perlo [12] for a Uniform$(-1, 1)$ parent.

Aside from the Gaussian family, the exponential location-scale family is also a remarkable member of the Koopman–Darmois–Pitman exponential family. It is the only member of this family with support in a half-line $[\lambda, \infty)$ and with a pair of sufficient statistics to estimate (λ, δ), namely $(X_{1:n}, \overline{X}_n - X_{1:n})$, where $X_{k:n}$, $k = 1, \ldots, n$, denotes the k-th order statistic of the random sample $(X_1, \ldots, X_n)$.

Moreover, the spacings $X_{k:n} - X_{k-1:n} \sim$ Exponential$\big(0, \frac{1}{n+1-k}\big)$, $k = 1, \ldots, n$, are independent, with the convention $X_{0:n} = \lambda$. This property characterizes the exponential distribution, and it is a consequence of Pexider's [13] functional equation $f(x+y) = \phi_1(x)\,\phi_2(y)$, which is an extension of Cauchy's functional equation, with the solution $f(x) = k_1\,\mathrm{e}^{k_2 x}$.

Using the independence of the spacings of the exponential model, we shall also obtain in Section 4 the probability density function of some externally and internally Studentized statistics, according to David's [14] definition, for inferences on the exponential location parameter. The independence of spacings is further used to compare the location parameters of two exponential populations in Subsection 4.4, and also to establish an ANalysis Of Spacings (ANOSp), similar to a one-way ANOVA in Gaussian populations, which is presented in Subsection 4.5. A Satterthwaite [15] approximation for the case of unequal dispersions is given in Subsection 4.6.

Finally, Section 5 summarizes the main findings regarding approximate solutions for non-Gaussian symmetric populations and for a very asymmetric exponential population, either for exact solutions resulting from the independence of spacings or for approximate solutions for inferences on the comparison of $k \geq 2$ location parameters without assuming equal dispersions.

2. Joint Probability Density Function of $(\overline{X}_n, SS_n)$ and Probability Density Function of T_{n-1}

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random sample of size n from a population X. If $X \sim$ Gaussian(μ, σ), then from the independence of $\overline{X}_n$ and SS_n, the joint probability density function of $(\overline{X}_n - \mu, SS_n)$ is

$$f_{(\overline{X}_n-\mu, SS_n)}(w,s) = f_{\overline{X}_n-\mu}(w) f_{SS_n}(s) = \frac{1}{\sqrt{2\pi}\,(\sigma/\sqrt{n})}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{w}{\sigma/\sqrt{n}}\right)^2}\,\frac{s^{\frac{n-1}{2}-1}\,\mathrm{e}^{-\frac{s}{2\sigma^2}}}{(2\sigma^2)^{\frac{n-1}{2}}\,\Gamma\big(\frac{n-1}{2}\big)}\,\mathbb{I}_{\mathbb{R}\times\mathbb{R}^+}(w,s)\,, \tag{4}$$

where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x$, $\alpha > 0$, is Euler's Gamma function.

2.1. *Joint Probability Density Function of* $(\overline{X}_2, SS_2)$

The independence of $\overline{X}_n$ and SS_n does not hold for a non-Gaussian population X. However, for samples of size $n = 2$, it is easy to obtain the probability density function of Student's statistic T_1.

Since

$$\overline{X}_2 = \frac{X_1 + X_2}{2} \quad \text{and} \quad SS_2 = \frac{(X_1 - X_2)^2}{2},$$

it follows that if $X_1 \geq X_2$, then

$$X_1 = \overline{X}_2 + \sqrt{\frac{SS_2}{2}} \quad \text{and} \quad X_2 = \overline{X}_2 - \sqrt{\frac{SS_2}{2}},$$

with a similar result also holding true if $X_1 < X_2$, namely $X_1 = \overline{X}_2 - \sqrt{\frac{SS_2}{2}}$ and $X_2 = \overline{X}_2 + \sqrt{\frac{SS_2}{2}}$. Therefore, if we denote f_X the probability density function of X, as the absolute value for the Jacobian determinant of both transformations is $|J| = \frac{1}{\sqrt{2SS_2}}$, then the joint probability density function of $(\overline{X}_2, SS_2)$ is

$$\begin{aligned} f_{(\overline{X}_2,SS_2)}(w,s) &= f_{X_1}\left(w+\sqrt{\tfrac{s}{2}}\right) f_{X_2}\left(w-\sqrt{\tfrac{s}{2}}\right)\frac{1}{\sqrt{2s}} + f_{X_1}\left(w-\sqrt{\tfrac{s}{2}}\right) f_{X_2}\left(w+\sqrt{\tfrac{s}{2}}\right)\frac{1}{\sqrt{2s}} \\ &= \sqrt{\frac{2}{s}}\, f_X\left(w+\sqrt{\tfrac{s}{2}}\right) f_X\left(w-\sqrt{\tfrac{s}{2}}\right) \\ &= \sqrt{\frac{2}{s}}\, f_X\left(w+\tfrac{1}{\sqrt{2}}\sqrt{s}\right) f_X\left(w-\tfrac{1}{\sqrt{2}}\sqrt{s}\right)\mathbb{I}_{\mathcal{S}}(s,w), \end{aligned} \tag{5}$$

where the support

$$\mathcal{S} = \begin{cases} \mathbb{R}\times\mathbb{R}^+ & \text{, if the support of } X \text{ is } \mathbb{R} \\ \mathbb{R}^+\times(0,2w^2) & \text{, if the support of } X \text{ is } \mathbb{R}^+ \\ [a,b]\times\left(0,2\min\{(w-a)^2,(b-w)^2\}\right) & \text{, if the support of } X \text{ is [a,b]} \end{cases}. \tag{6}$$

Note that if in (5), $X \sim \text{Gaussian}(0,\sigma)$, then

$$f_{(\overline{X}_2,SS_2)}(w,s) = \sqrt{\frac{2}{s}}\,\frac{\mathrm{e}^{-\frac{s+2w^2}{2\sigma^2}}}{2\pi\sigma^2} = \frac{\mathrm{e}^{-\frac{1}{2}\left(\frac{w}{\sigma/\sqrt{2}}\right)^2}}{\sqrt{2\pi}(\sigma/\sqrt{2})}\,\frac{s^{\frac{1}{2}-1}\mathrm{e}^{-\frac{s}{2\sigma^2}}}{(2\sigma^2)^{\frac{1}{2}}\Gamma\left(\frac{1}{2}\right)}\,\mathbb{I}_{\mathbb{R}\times\mathbb{R}^+}(w,s),$$

i.e., the same expression as the one given by (4) for $n = 2$ because $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

As for the coefficients of $\sqrt{s}$ in the arguments of the functions in (5), they satisfy the conditions

$$\frac{1}{\sqrt{2}} + \left(-\frac{1}{\sqrt{2}}\right) = 0 \quad \text{and} \quad \left(\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2 = 1.$$

2.2. Probability Density Function of Student's Statistic T_1

To obtain the probability density function of Student's statistic T_1, let

$$T_1 = \frac{\sqrt{2}\,\overline{X}_2}{\sqrt{SS_2}} \quad \text{and} \quad U = \sqrt{SS_2},$$

from which the inverse transformation is $\overline{X}_2 = \frac{UT_1}{\sqrt{2}}$ and $SS_2 = U^2$, with $|J| = \sqrt{2}U^2$. Hence, the joint probability density function of (T_1, U) is

$$f_{(T_1,U)}(t,u) = 2u\, f_X\left(\frac{u(t+1)}{\sqrt{2}}\right) f_X\left(\frac{u(t-1)}{\sqrt{2}}\right),$$

and therefore, if X has support $\mathbb{R}$, the probability density function of T_1 is given by

$$f_{T_1}(t) = 2\int_0^\infty u\, f_X\left(\frac{u(t+1)}{\sqrt{2}}\right) f_X\left(\frac{u(t-1)}{\sqrt{2}}\right) \mathrm{d}u. \tag{7}$$

Some simple examples for symmetric parent distributions are the following:

- If $X \sim \text{Gaussian}(0,1)$,
$$f_{T_1}(t) = \frac{1}{\pi(1+t^2)}\,\mathbb{I}_{\mathbb{R}}(t)\,,$$
i.e., $T_1 \sim \text{Cauchy}(0,1)$ (note that a standard Cauchy random variable is Student t-distributed with one degree of freedom).
- If $X \sim \text{Uniform}(-1,1)$,
$$f_{T_1}(t) = \frac{1}{2(1+|t|)^2}\,\mathbb{I}_{\mathbb{R}}(t)\,.$$
- If $X \sim \text{Beta}(2,2;-1,1)$,
$$f_{T_1}(t) = \frac{3}{8}\,\frac{1+4|t|+t^2}{(1+|t|)^4}\mathbb{I}_{\mathbb{R}}(t)\,.$$
- If $X \sim \text{Laplace}(0,1)$,
$$f_{T_1}(t) = \begin{cases} 1/4 & |t| \le 1 \\ 1/(4t^2) & |t| > 1 \end{cases}\,.$$
- If $X \sim \text{Cauchy}(0,1)$,
$$f_{T_1}(t) = \frac{1}{2\pi^2 t}\ln\left(\frac{t+1}{t-1}\right)^2\,\mathbb{I}_{\mathbb{R}-\{-1,1\}}(t)\,.$$

(Note that a random variable X has a symmetric distribution around a parameter $\theta \in \mathbb{R}$ if its cumulative distribution function F_X satisfies $F_X(\theta - x) = F_X(\theta + x)$, $x \in \mathbb{R}$. In particular, if X is an absolutely continuous random variable, its distribution is symmetric around θ if $f_X(\theta - x) = f_X(\theta + x)$, $x \in \mathbb{R}$.)

The graphics in Figure 1 show that the probability density function of T_1 with a non-Gaussian parent can be quite different from that of t_1, i.e., with a Gaussian parent. Note that with a $\text{Uniform}(-1,1)$ parent and a $\text{Beta}(2,2;-1,1)$ parent, the corresponding densities of T_1 are unimodal and have heavier tails than those of t_1 (see Figure 2). In the case of a Cauchy parent, which is known for having very heavy tails, T_1 has an antimodal probability density function. The observation that heavier tails of the underlying parent distribution result in less heavy tails for the distribution of Student's T-statistic, and vice-versa, was proved by van Zwet [16,17]. This indicates that the tail weight of the distribution is very important to determine the conservative or liberal behavior of Student's statistic.

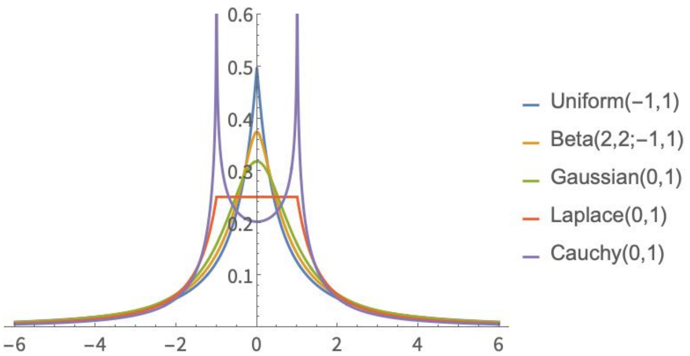

Figure 1. Probability density function of T_1 with symmetric parent distributions.

2.3. *Joint Probability Density Function of* $(\overline{X}_n, SS_n)$

Let $\mathbf{X} = (X_1, \ldots, X_{n+1})$, $n \ge 2$, be a random sample from the parent distribution X. From $(\overline{X}_n, SS_n)$ and X_{n+1} being independent, the joint probability density function of $(\overline{X}_n, SS_n, X_{n+1})$ is

$$f_{(\overline{X}_n,SS_n,X_{n+1})}(w,s,x) = f_{(\overline{X}_n,SS_n)}(w,s)\, f_{X_{n+1}}(x) = f_{(\overline{X}_n,SS_n)}(w,s)\, f_X(x)\,.$$

On the other hand, as

$$\overline{X}_{n+1} = \frac{n}{n+1}\overline{X}_n + \frac{X_{n+1}}{n+1} \quad \text{and} \quad SS_{n+1} = SS_n + \frac{n}{n+1}(\overline{X}_n - X_{n+1})^2\,,$$

using the auxiliary random variable $Y = \overline{X}_n - X_{n+1}$, the inverse transformation is

$$\overline{X}_n = \overline{X}_{n+1} + \frac{Y}{n+1}\,,\; SS_n = SS_{n+1} - \frac{n}{n+1}Y^2 \;\text{ and }\; X_{n+1} = \overline{X}_{n+1} - \frac{n}{n+1}Y\,,$$

with $|J| = 1$. Hence, the joint probability density function of $(\overline{X}_{n+1}, SS_{n+1}, Y)$ is

$$f_{(\overline{X}_{n+1},SS_{n+1},Y)}(w,s,y) = f_{(\overline{X}_n,SS_n)}\left(w + \frac{y}{n+1}, s - \frac{ny^2}{n+1}\right) f_X\left(w - \frac{ny}{n+1}\right),$$

and therefore, the joint probability density function of $(\overline{X}_{n+1}, SS_{n+1})$ is given by

$$f_{(\overline{X}_{n+1},SS_{n+1})}(w,s) = \int_{-\sqrt{\frac{n+1}{n}s}}^{\sqrt{\frac{n+1}{n}s}} f_{(\overline{X}_n,SS_n)}\left(w + \frac{y}{n+1}, s - \frac{ny^2}{n+1}\right) f_X\left(w - \frac{ny}{n+1}\right) \mathrm{d}y\,. \tag{8}$$

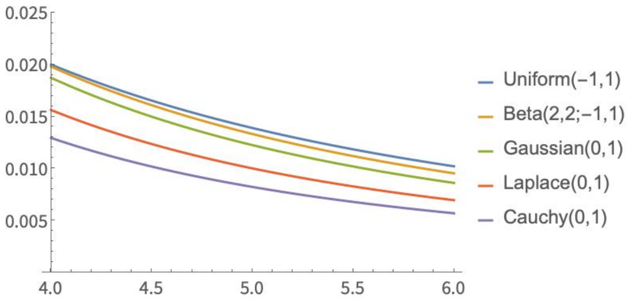

Figure 2. Right tail of the probability density function of T_1 for the symmetric parent distributions in Figure 1.

We emphasize that in integral (8), it is assumed that X has support $\mathbb{R}$, which we shall continue to assume for what follows, but if this is not the case, then the integration limits should be defined according to its support (see (6)).

In making the replacement $v = \frac{y}{\sqrt{\frac{n+1}{n}s}}$ in (8), it follows that

$$f_{(\overline{X}_{n+1},SS_{n+1})}(w,s) = \int_{-1}^{1} \sqrt{\tfrac{n+1}{n}s}\, f_{(\overline{X}_n,SS_n)}\left(w + v\sqrt{\tfrac{s}{n(n+1)}}, s(1-v^2)\right) f_X\left(w - v\sqrt{\tfrac{ns}{n+1}}\right) \mathrm{d}v. \tag{9}$$

For example, if $n+1 = 3$, then in using Formulas (9) and (5),

$$\begin{aligned} f_{(\overline{X}_3,SS_3)}(w,s) &= \int_{-1}^{1} \sqrt{\frac{3}{2}s}\, f_{(\overline{X}_2,SS_2)}\left(w + v\sqrt{\tfrac{s}{6}}, s(1-v^2)\right) f_X\left(w - v\sqrt{\tfrac{2}{3}s}\right) \mathrm{d}v \\ &= \sqrt{3}\int_{-1}^{1} \frac{1}{\sqrt{1-v^2}}\, f_X\left(w + \left(v + \sqrt{3(1-v^2)}\,\right)\sqrt{\tfrac{s}{6}}\right) \\ &\qquad \times f_X\left(w + \left(v - \sqrt{3(1-v^2)}\,\right)\sqrt{\tfrac{s}{6}}\right) f_X\left(w - 2v\sqrt{\tfrac{s}{6}}\right) \mathrm{d}v \\ &= \sqrt{3}\int_{-1}^{1} \frac{1}{\sqrt{1-v^2}}\, f_X\left(w + \tfrac{v+\sqrt{3(1-v^2)}}{\sqrt{6}}\sqrt{s}\right) \\ &\qquad \times f_X\left(w + \tfrac{v-\sqrt{3(1-v^2)}}{\sqrt{6}}\sqrt{s}\right) f_X\left(w - \tfrac{2v}{\sqrt{6}}\sqrt{s}\right) \mathrm{d}v\,. \end{aligned} \tag{10}$$

Notice that the coefficients of $\sqrt{s}$ in the arguments of the functions in the last integrand in (10) satisfy the following two conditions:

$$\frac{v+\sqrt{3(1-v^2)}}{\sqrt{6}}+\frac{v-\sqrt{3(1-v^2)}}{\sqrt{6}}-\frac{2v}{\sqrt{6}}=0$$

and

$$\left(\frac{v+\sqrt{3(1-v^2)}}{\sqrt{6}}\right)^2+\left(\frac{v-\sqrt{3(1-v^2)}}{\sqrt{6}}\right)^2+\left(-\frac{2v}{\sqrt{6}}\right)^2=1\,.$$

2.4. The Gaussian Case

In applying (10), in particular, to the case $X\sim \text{Gaussian}(0,\sigma)$, for $w\in\mathbb{R}$ and $s>0$,

$$\begin{aligned}
f_{(\overline{X}_3,SS_3)}(w,s) &= \sqrt{3}\int_{-1}^{1}\frac{1}{\sqrt{1-v^2}}\frac{\mathrm{e}^{-\frac{s+3w^2}{2\sigma^2}}}{2\sqrt{2}\,\pi^{3/2}\sigma^3}\mathrm{d}v\\
&= \sqrt{3}\frac{\mathrm{e}^{-\frac{s+3w^2}{2\sigma^2}}}{2\sqrt{2}\,\pi^{3/2}\sigma^3}\int_{-1}^{1}\frac{1}{\sqrt{1-v^2}}\mathrm{d}v\\
&= \frac{1}{\sqrt{2\pi}\,\frac{\sigma}{\sqrt{3}}}\mathrm{e}^{-\frac{3w^2}{2\sigma^2}}\mathbb{I}_{\mathbb{R}}(w)\,\frac{\mathrm{e}^{-\frac{s}{2\sigma^2}}}{2\sigma^2}\mathbb{I}_{\mathbb{R}^+}(s)\\
&\approx f_X\left(w+\sqrt{\tfrac{s}{2}}\right)f_X(w)\,f_X\left(w-\sqrt{\tfrac{s}{2}}\right),
\end{aligned}\tag{11}$$

where $\left\{\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}}\right\}$ is a set of equidistant points such that their sum is 0 and the sum of their squares is 1.

As for the joint probability density function of $(\overline{X}_4,SS_4)$, using the recurrence Formula (9), we obtain

$$f_{(\overline{X}_4,SS_4)}(w,s)=\sqrt{\frac{4s}{3}}\int_{-1}^{1}f_{(\overline{X}_3,SS_3)}\left(w+\sqrt{\tfrac{s}{12}}\,v,(1-v^2)s\right)f_X\left(w-\sqrt{\tfrac{3s}{4}}\,v\right)\mathrm{d}v\,,\tag{12}$$

and in view of what was established in (11), the joint probability density function (12) can be approximated using the product of f_X computed at the points $w+\alpha_{i,4}\sqrt{s}$, $i=1,2,3,4$, i.e.,

$$f_{(\overline{X}_4,SS_4)}(w,s)\approx K\sqrt{s}\prod_{i=1}^{4}f_X\left(w+\alpha_{i,4}\sqrt{s}\right),\tag{13}$$

with K as a norming constant, provided that $\sum_{i=1}^{4}\alpha_{i,4}=0$ and $\sum_{i=1}^{4}\alpha_{i,4}^2=1$.

In determining the points $w+\alpha_{i,4}\sqrt{s}$, since the set $A=\{\alpha_{1,4},\alpha_{2,4},\alpha_{3,4},\alpha_{3,4}\}$ is symmetric, for its elements to fulfill the first condition, they must satisfy $a_{1,4}=-a_{4,4}$ and $a_{2,4}=-a_{3,4}$. As the points are also equidistant, i.e., $\alpha_{i,4}-\alpha_{i-1,4}=r$, $i=2,3,4$, they form an arithmetic progression with common distance r, and thus, $A=\left\{\frac{3r}{2},\frac{r}{2},-\frac{r}{2},-\frac{3r}{2}\right\}$. In order to satisfy the second condition, we must also have $r^2=\frac{1}{5}$, and hence,

$$A=\{\alpha_{1,4},\alpha_{2,4},\alpha_{3,4},\alpha_{3,4}\}=\left\{\frac{3}{\sqrt{20}},\frac{1}{\sqrt{20}},-\frac{1}{\sqrt{20}},-\frac{3}{\sqrt{20}}\right\}.$$

Therefore, from (13), and for $w\in\mathbb{R}$ and $s>0$, we obtain

$$f_{(\overline{X}_4,SS_4)}(w,s)\approx \mathrm{e}^{-\frac{4w^2}{2\sigma^2}}\sqrt{s}\,\mathrm{e}^{-\frac{s}{2\sigma^2}}\propto f_{\overline{X}_4}(w)\,f_{SS_4}(s).\tag{14}$$

More generally, we have the following:

- If $n=2k+1$, then

$$\{\alpha_{1,n},\ldots,\alpha_{n,n}\}=\left\{\frac{n-1}{2}r,\left(\frac{n-1}{2}-1\right)r,\ldots,r,0,-r,\ldots,-\left(\frac{n-1}{2}-1\right)r,-\frac{n-1}{2}r\right\},$$

and from $2\,\frac{\frac{n-1}{2}\,\frac{n+1}{2}\,n}{6}\,r^2 = 1$, we obtain $r^2 = \frac{12}{n(n^2-1)}$.

- If $n = 2k$, then

$$\{\alpha_{1,n}, \ldots, \alpha_{n,n}\} = \left\{\frac{2n-1}{2}r, \frac{2n-3}{2}r, \ldots, \frac{r}{2}, -\frac{r}{2}, \ldots, -\frac{2n-3}{2}r, -\frac{2n-1}{2}r\right\},$$

and from $2\sum_{k=0}^{n/2}\left(\frac{2k-1}{2}\right)^2 r^2 = 1$, we obtain $r^2 = \frac{12}{n(n^2-1)}$.

It follows that $\alpha_{1,n} = \sqrt{\frac{3(n-1)}{n(n+1)}}$, and since $\alpha_{i,n} = \left(\frac{n-1}{2} - i + 1\right)r$, $i = 1, \ldots, n$, we have

$$\alpha_{i,n} = (n+1-2i)\sqrt{\frac{3}{n(n^2-1)}}\,, \quad i = 1, \ldots, n\,.$$

In particular, $\alpha_{i,n-1} = (n-2i)\sqrt{\frac{3}{(n-1)n(n-2)}}$, and hence,

$$\sqrt{\frac{n-2}{n+1}}\,\alpha_{i,n-1} = (n-2i)\sqrt{\frac{3}{n(n^2-1)}}\,.$$

Therefore,

$$\alpha_{i,n} - \sqrt{\frac{n-2}{n+1}}\,\alpha_{i,n-1} = \sqrt{\frac{3}{n(n^2-1)}}\,\left(=\frac{r}{2}\right).$$

For example, for $n = 2$ and $n = 3$ we obtained the values

$$n = 2: \quad \alpha_{1,2} = \frac{1}{\sqrt{2}}, \quad \alpha_{2,2} = -\frac{1}{\sqrt{2}};$$
$$n = 3: \quad \alpha_{1,3} = \frac{1}{\sqrt{2}}, \quad \alpha_{2,3} = 0, \quad \alpha_{3,3} = -\frac{1}{\sqrt{2}}.$$

As a recurrence formula,

$$\alpha_{i,n} = \sqrt{\frac{3}{n(n^2-1)}} + \sqrt{\frac{n-2}{n+1}}\,\alpha_{i,n-1}, \quad i = 1, \ldots, n-1, \qquad \alpha_{n,n} = -\sqrt{\frac{3(n-1)}{n(n+1)}}. \tag{15}$$

By expressing the multiplier $\sqrt{\frac{n-2}{n+1}}$ as $\sqrt{\frac{n-2}{n+1}} = \sqrt{1-\xi_n^2}$, we obtain $\xi_n = \sqrt{\frac{3}{n+1}}$.

3. Symmetry and Studentization

3.1. Symmetric Random Variables

There are useful results that enable the identification or characterization of symmetric random variables, such as the following:

(i) If X and Y are independent random variables, then $W = XY$ has a characteristic function $\varphi_W(t) = \mathbb{E}\left(e^{itXY}\right) = \int_{-\infty}^{\infty}\mathbb{E}\left(e^{itXy}\right)dF_Y(y) = \int_{-\infty}^{\infty}\varphi_X(ty)\,dF_Y(y)$.

(ii) If $X \sim \text{Bernoulli}\left(\frac{1}{2}\right)$, then $B = 2X - 1$ has a probability mass function $\mathbb{P}(B = -1) = \mathbb{P}(B = 1) = \frac{1}{2}$, and its characteristic function is $\varphi_B(t) = \mathbb{E}\left(e^{itB}\right) = \frac{1}{2}e^{-it} + \frac{1}{2}e^{it} = \cos t$.

(iii) If X and B (B as defined in (ii)) are independent, the characteristic function of $W = BX$ is $\varphi_W(t) = \int_{-\infty}^{\infty}\cos(tx)\,dF_X(x)$.

(iv) A random variable is symmetric if and only if its characteristic function is $\varphi_X(t) = \int_{-\infty}^{\infty}\cos(tx)\,dF_X(x)$.

From the above, a random variable X is symmetric if and only if $X \stackrel{d}{=} XB$, with X and B being independent. Therefore, if X and Y are independent random variables and

X is symmetric, then XY is also symmetric. In particular, if X is symmetric, then $\overline{X}_n$ is symmetric, which leads to T_{n-1} being symmetric as well.

3.2. An Approximate Joint Probability Density Function of $(\overline{X}_n, SS_n)$ with a Symmetric Parent Distribution

Using the Gaussian case as a guideline, for $n = 3$, we assume that f_X is a continuous and smooth probability density function of a symmetric random variable X, in the sense that there is a value $\xi_3(f_X) \in (-1, 1)$ such that, from the mean value theorem, the integral in (10) can be computed as

$$\begin{aligned} f_{(\overline{X}_3, SS_3)}(w,s) &= \frac{1}{\sqrt{1-\xi_3^2}}\, f_X\left(w + \frac{\xi_3 + \sqrt{3(1-\xi_3^2)}}{\sqrt{6}}\sqrt{s}\right) \\ &\qquad \times f_X\left(w - \frac{\xi_3 + \sqrt{3(1-\xi_3^2)}}{\sqrt{6}}\sqrt{s}\right) f_X\left(w - \frac{2\xi_3}{\sqrt{6}}\sqrt{s}\right) \\ &= \frac{1}{\sqrt{1-\xi_3^2}} \prod_{i=1}^{3} f_X(w + \alpha_{i,3}(\xi_3)\sqrt{s}), \end{aligned} \tag{16}$$

where $\alpha_{1,3}(\xi_3) = \frac{\xi_3 + \sqrt{3(1-\xi_3^2)}}{\sqrt{6}}$, $\alpha_{2,3}(\xi_3) = \frac{-\xi_3 - \sqrt{3(1-\xi_3^2)}}{\sqrt{6}}$, and $\alpha_{2,3}(\xi_3) = -\frac{3\,\xi_3}{\sqrt{6}}$, so that $\sum_{i=1}^{3} \alpha_{i,3}(\xi_3) = 0$ and $\sum_{i=1}^{3} \alpha_{i,3}^2(\xi_3) = 1$.

More generally, we investigate the consequences of a smoothness hypothesis, which is based on the Gaussian case.

Smoothness Hypothesis 1. Let $X_1, X_2, \ldots$ be independent replicas of a random variable X, with a smooth probability density function f_X that allows the integral mean value theorem to be used. Therefore, for $\nu = 2, \ldots, n$, there are values $\xi_\nu(f_X) \in (-1, 1)$ and $\alpha_{i,\nu}(\xi_\nu)$ of the form

$$\alpha_{i,\nu} = \frac{\xi_\nu}{\sqrt{\nu(\nu-1)}} + \alpha_{i,\nu-1}\sqrt{1-\xi_\nu^2}\,,\; i = 1, \ldots, \nu-1, \quad \text{and} \quad \alpha_{\nu,\nu} = -\xi_\nu \sqrt{\frac{\nu-1}{\nu}}\,,$$

satisfying the conditions $\sum_{i=1}^{\nu} \alpha_{i,\nu}(\xi_\nu) = 0$ and $\sum_{i=1}^{\nu} \alpha_{i,\nu}^2(\xi_\nu) = 1$ such that the approximation

$$f_{(\overline{X}_\nu, SS_\nu)}(w,s) \approx K\, s^{(\nu-3)/2} \prod_{i=1}^{\nu} f_X(w + \alpha_{i,\nu}\sqrt{s})\, \mathbb{I}_{\mathcal{S}}(w,s)\,, \tag{17}$$

where K is a norming constant, is valid.

Then, the integral mean value theorem also holds for $f_{(\overline{X}_{n+1}, SS_{n+1})}$, i.e., there are values $\xi_{n+1}(f_X) \in (-1, 1)$ and $\alpha_{i,n+1}(\xi_{n+1})$ such that the approximation

$$f_{(\overline{X}_{n+1}, SS_{n+1})}(w,s) \approx K\, s^{(n+1-3)/2} \prod_{i=1}^{n+1} f_X(w + \alpha_{i,n+1}(\xi_{n+1})\sqrt{s})\, \mathbb{I}_{\mathcal{S}}(w,s) \tag{18}$$

is also valid, with $(\overline{X}_{n+1}, SS_{n+1})$ having support $\mathcal{S}$ defined in (6), and $\alpha_{i,n+1} = \frac{\xi_{n+1}}{\sqrt{n(n+1)}} + \alpha_{i,n}\sqrt{1-\xi_{n+1}^2}$, $i = 1, \ldots, n$, and $\alpha_{n+1,n+1} = -\xi_{n+1}\sqrt{\frac{n}{n+1}}$.

If we further assume that X is a symmetric random variable, which implies that the points $\alpha_{1,n}, \ldots, \alpha_{n,n}$ form an arithmetic progression with common distance $r_n = -\frac{2\sqrt{3}}{\sqrt{n(n^2-1)}}$ and $\alpha_{i,n} = -\alpha_{n+1-i,n}$, $i = 1, \ldots, n$, then $\xi_n = \sqrt{\frac{3}{n+1}}$ and

$$\alpha_{i,n} = (n+1-2i)\sqrt{\frac{3}{n(n^2-1)}}\,,\; i = 1, \ldots, n. \tag{19}$$

Equation (19) is easily obtained. In fact, if $n = 2k+1$, then $\alpha_{1,n} = k|r_n|$. Moreover, the abcissas of the positive points $\alpha_{i,n}$ are $|r|, 2|r|, \ldots, k|r|$, and since $\sum_{j=1}^{k}(jr)^2 = \frac{1}{2}$, this implies that $|r| = 2\sqrt{\frac{3}{n(n^2-1)}}$. For $n = 2k$, the results are similar, with the abcissas of the positive $\alpha_{i,n}$ now being $\frac{|r|}{2}, \frac{3|r|}{2}, \ldots, (k-\frac{1}{2})|r|$.

3.3. An Approximate Expression for the Probability Density Function of T_{n-1} with a Smooth Symmetric Parent

Considering the transformation

$$T_{n-1} = \frac{\sqrt{n(n-1)}\,\overline{X}_n}{\sqrt{SS_n}} \quad \text{and} \quad U = \sqrt{SS_n}\,,$$

from which $\overline{X}_n = \frac{UT_{n-1}}{\sqrt{n(n-1)}}$ and $SS_n = U^2$, with $|J| = \frac{2U^2}{\sqrt{n(n-1)}}$, we obtain

$$f_{(T_{n-1},U)}(t,u) = \frac{2u^2}{\sqrt{n(n-1)}} f_{(\overline{X}_n, SS_n)}\left(\frac{ut}{\sqrt{n(n-1)}}, u^2\right),$$

and therefore, for $t \in \mathbb{R}$, and X with support $\mathbb{R}$,

$$f_{T_{n-1}}(t) = \frac{2}{\sqrt{n(n-1)}} \int_0^\infty u^2 f_{(\overline{X}_n, SS_n)}\left(\frac{ut}{\sqrt{n(n-1)}}, u^2\right) \mathrm{d}u\,.$$

Using approximation (18), and denoting $\alpha_{i,n}(\xi_n) = \alpha_{i,n}$, it follows that

$$f_{T_{n-1}}(t) \approx K \int_0^\infty u^{n-1} \prod_{i=1}^{n} f_X\left(\left(\frac{t}{\sqrt{n(n-1)}} + \alpha_{i,n}\right)u\right) \mathrm{d}u\,. \tag{20}$$

For example, if $X \sim \text{Gaussian}(0,1)$, then

$$\begin{aligned} f_{t_{n-1}}(t) &\approx K \int_0^\infty u^{n-1} \prod_{i=1}^{n} \exp\left\{-\frac{1}{2}\left[\left(\frac{t}{\sqrt{n(n-1)}} + \alpha_{i,n}\right)u\right]^2\right\} \mathrm{d}u = \\ &= K \int_0^\infty u^{n-1} \exp\left[-\frac{u^2}{2}\left(\frac{t^2}{n-1}+1\right)\right] \mathrm{d}u = K \frac{1}{\left(1+\frac{t^2}{n-1}\right)^{n/2}}, \end{aligned}$$

as expected.

If $X \sim \text{Laplace}(0,1)$, where $f_X(x) = \frac{1}{2}\mathrm{e}^{-|x|}$, then

$$\begin{aligned} f_{T_{n-1}}(t) &\approx K \int_0^\infty u^{n-1} \prod_{i=1}^{n} \exp\left[-\left|\left(\frac{t}{\sqrt{n(n-1)}} + \alpha_{i,n}\right)u\right|\right] \mathrm{d}u = \\ &= K \int_0^\infty u^{n-1} \exp\left(-u \sum_{i=1}^{n} \left|\frac{t}{\sqrt{n(n-1)}} + \alpha_{i,n}\right|\right) \mathrm{d}u\,, \end{aligned}$$

and therefore,

$$f_{T_{n-1}}(t) \approx K \left(\sum_{i=1}^{n} \left|\frac{t}{\sqrt{n(n-1)}} + \alpha_{i,n}\right|\right)^{-n} \mathbb{I}_{\mathbb{R}}(t)\,.$$

In particular, if $n = 3$,

$$f_{T_2}(t) \approx K \left(\left|\frac{t}{\sqrt{6}} + \frac{1}{\sqrt{2}}\right| + \frac{|t|}{\sqrt{6}} + \left|\frac{t}{\sqrt{6}} - \frac{1}{\sqrt{2}}\right|\right)^{-3} \mathbb{I}_{\mathbb{R}}(t)\,. \tag{21}$$

Computing the norming constant so that the right-hand side of (21) is transformed into a probability density function, we obtain

$$f_{T_2}(t) \approx \frac{9\sqrt{6}}{19}\left(\left|\frac{t}{\sqrt{6}}+\frac{1}{\sqrt{2}}\right|+\frac{|t|}{\sqrt{6}}+\left|\frac{t}{\sqrt{6}}-\frac{1}{\sqrt{2}}\right|\right)^{-3}\mathbb{I}_{\mathbb{R}}(t)\,.$$

In Figure 3, the approximate probability density function of T_2 with a Laplace$(0,1)$ parent is represented.

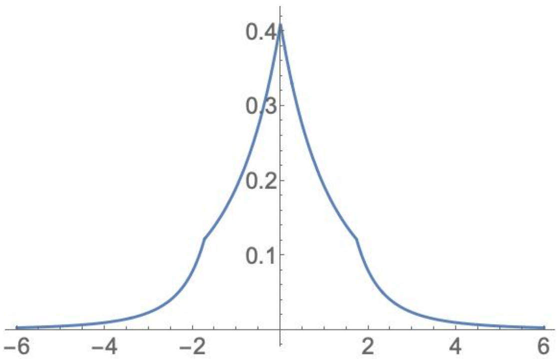

Figure 3. Approximate probability density function of T_2 with a Laplace$(0,1)$ parent.

For the non-Gaussian parent case, to assess how well approximation (20) works, we shall compare the exact probability density function of T_2 with a Uniform$(-1,1)$ parent, given by Perlo [12], namely

$$f_{T_2}(t) = \begin{cases} \frac{\sqrt{3}\left(1-\frac{9t^2}{4-t^2}\right)}{2(4-t^2)\sqrt{1-t^2}} + \frac{3\sqrt{3}(t^2+2)}{(4-t^2)^{5/2}}\operatorname{arctanh}\left(\sqrt{\frac{1-t^2}{4-t^2}}\right) & , |t| \le \frac{1}{2} \\ \frac{9\left(\frac{1}{|t|+1)}+\frac{3|t|}{t^2-4}\right)}{4(|t|+1)(4-t^2)} + \frac{3\sqrt{3}(t^2+2)}{(t^2-4)^{5/2}}\arctan\left(\frac{\sqrt{t^2-4}}{\sqrt{3}(t+2)}\right) & , |t| > \frac{1}{2} \end{cases}, \tag{22}$$

with the approximation using (20).

As we are now dealing with a parent distribution with a limited support, the integration limits for (20) must be defined according to (6). In this case, the condition

$$-\sqrt{6} < ut < \sqrt{6} \,\wedge\, 0 < u^2 < 2\min\left\{\left(\frac{ut}{\sqrt{6}}+1\right)^2, \left(1-\frac{ut}{\sqrt{6}}\right)^2\right\}$$

must be satisfied, or equivalently,

$$-\sqrt{6} < ut < \sqrt{6} \,\wedge\, 0 < u < \min\left\{\sqrt{2}\left(\frac{ut}{\sqrt{6}}+1\right), \sqrt{2}\left(1-\frac{ut}{\sqrt{6}}\right)\right\}.$$

Hence, in denoting $a = \min\left\{\sqrt{2}\left(\frac{ut}{\sqrt{6}}+1\right), \sqrt{2}\left(1-\frac{ut}{\sqrt{6}}\right)\right\}$, it follows that

$$f_{T_2}(t) \approx K\int_0^a u^2\left(\frac{1}{2}\right)^3 \mathrm{d}u \propto \int_0^a u^2 \mathrm{d}u = \frac{2\sqrt{6}}{(\sqrt{3}+|t|)^3}\mathbb{I}_{\mathbb{R}}(t)\,. \tag{23}$$

Computing the norming constant K so that the function in the right-hand side of (23) becomes a probability density function, we obtain

$$f_{T_2}(t) \approx \frac{3}{(\sqrt{3}+|t|)^3}\mathbb{I}_{\mathbb{R}}(t)\,. \tag{24}$$

In Figure 4, the exact and approximate probability density functions of T_2 with a Uniform$(-1,1)$ parent, defined in (22) and (24), respectively, are plotted together. As can be observed, the approximation is not good in the central region of the distributions, but it is quite good in the tails of the distributions, which is where it actually matters because inferences are made using the tails of the distributions.

Observe that Hendriks et al. [11] analyzed why the Uniform parent is the least favorable one, as far as approximations are concerned, for the probability density function of T_{n-1}.

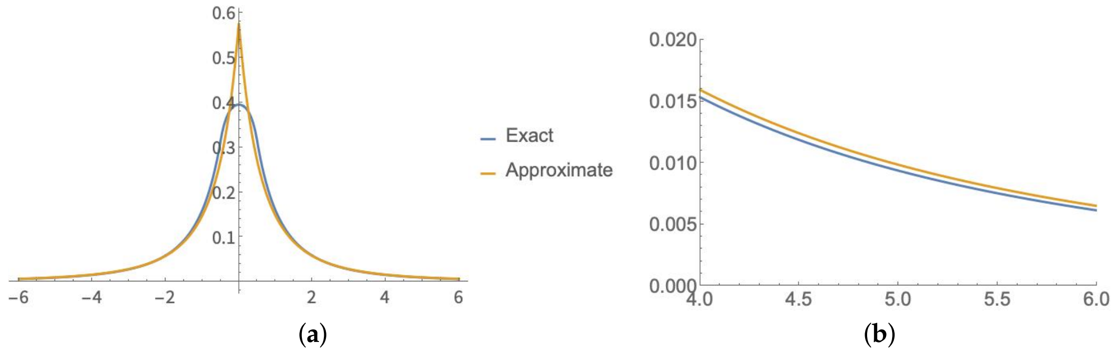

Figure 4. Probability density function of T_2 with a Uniform$(-1,1)$ parent: (**a**) Exact and approximate densities. (**b**) Zoom in of right tails.

4. Externally Studentized Statistics Using Spacings of an Exponential Parent

4.1. External Studentization Using the Maximum Likelihood Scale Estimator

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random sample of size n from $X \sim Exponential(\lambda, \delta)$, $\lambda \in \mathbb{R}$, $\delta > 0$, i.e., with probability density function

$$f_X(x) = \frac{1}{\delta}\,\mathrm{e}^{-\frac{x-\lambda}{\delta}}\,\mathbb{I}_{(\lambda,\infty)}(x)\,.$$

The maximum likelihood estimators $\widehat{\lambda} = X_{1:n} \sim \text{Exponential}\left(\lambda, \frac{\delta}{n}\right)$ and $\widehat{\delta} = \overline{X}_n - X_{1:n} = \frac{1}{n}\sum_{k=1}^{n}(X_{k:n} - X_{1:n}) = \frac{1}{n}\sum_{k=2}^{n}(n+1-k)(X_{k:n} - X_{k-1:n}) \sim \text{Gamma}\left(n-1, \frac{\delta}{n}\right)$ are independent due to the independence of the spacings $X_{k:n} - X_{k-1:n} \sim \text{Exponential}\left(0, \frac{\delta}{n+1-k}\right)$, $k = 1, \ldots, n$, with the usual convention $X_{0:n} = \lambda$.

Therefore, the externally Studentized statistic

$$T^*_{n-1} = \frac{X_{1:n} - \lambda}{\overline{X}_n - X_{1:n}}\,, \tag{25}$$

where $T^*_{n-1} \overset{d}{=} \frac{Z_{1:n}}{\overline{Z}_n - Z_{1:n}}$, with $(Z_1, \ldots, Z_n)$ being a random sample from $Z = \frac{X-\lambda}{\delta} \sim \text{Exponential}(0,1)$, has probability density function

$$\begin{aligned} f_{T^*_{n-1}}(t) &= \int_{-\infty}^{\infty} f_{Z_{1:n}}(tx) f_{\overline{Z}_n - Z_{1:n}}(x)|x|\mathrm{d}x \\ &= \int_0^{\infty} n\mathrm{e}^{-ntx}\frac{n^{n-1}x^{n-2}\mathrm{e}^{-nx}}{\Gamma(n-1)}x\,\mathrm{d}x \\ &= \frac{n^n}{\Gamma(n-1)}\int_0^{\infty} x^{n-1}\mathrm{e}^{-n(t+1)x}\mathrm{d}x \\ &= \frac{n-1}{(1+t)^n}\mathbb{I}_{(0,\infty)}(t)\,, \end{aligned}$$

and hence, $T^*_{n-1} \sim \text{Pareto}(n-1, 0)$.

4.2. External Studentization Using the Sample Range as a Dispersion Estimator

Another externally Studentized statistic that can be used to make inferences on the location parameter λ is

$$\tau_{n-1} = \frac{X_{1:n} - \lambda}{X_{n:n} - X_{1:n}}\,. \tag{26}$$

In noticing that $\tau_{n-1} \overset{d}{=} \frac{Z_{1:n}}{Z_{n:n}-Z_{1:n}}$ and considering that $Z_{n:n} - Z_{1:n} \overset{d}{=} Z_{n-1:n-1}$, the joint probability density function of $(Z_{1:n}, Z_{n:n} - Z_{1:n})$ is

$$f_{(Z_{1:n}, Z_{n:n}-Z_{1:n})}(x,y) = f_{Z_{1:n}}(x) f_{Z_{n:n}-Z_{1:n}}(y) = n\mathrm{e}^{-nx}(n-1)\left(1-\mathrm{e}^{-y}\right)^{n-2}\mathrm{e}^{-y}\mathbb{I}_{\mathbb{R}^+\times\mathbb{R}^+}(x,y)\,.$$

Using the transformation $S = Z_{1:n}$ and τ_{n-1}, with $|J| = \frac{S}{\tau_{n-1}^2}$, the joint probability density function of (S, τ_{n-1}) is defined as

$$f_{(S,\tau_{n-1})}(s,t) = n(n-1)\,\mathrm{e}^{-ns}\mathrm{e}^{-s/t}(1-\mathrm{e}^{-s/t})^{n-2}\,\frac{s}{t^2}\mathbb{I}_{\mathbb{R}^+\times\mathbb{R}^+}(s,t)\,.$$

Therefore, for $t > 0$,

$$\begin{aligned}
f_{\tau_{n-1}}(t) &= \frac{1}{t}\int_0^\infty ns\,\mathrm{e}^{-ns}\left[(n-1)\mathrm{e}^{-s/t}(1-\mathrm{e}^{-s/t})^{n-2}\frac{1}{t}\right]\mathrm{d}s \\
&= -\frac{1}{t}\int_0^\infty n(1-ns)\,\mathrm{e}^{-ns}(1-\mathrm{e}^{-s/t})^{n-1}\mathrm{d}s \\
&= -\frac{1}{t}\int_0^\infty n\,\mathrm{e}^{-ns}(1-\mathrm{e}^{-s/t})^{n-1}\mathrm{d}s + \frac{1}{t}\int_0^\infty n^2 s\,\mathrm{e}^{-ns}(1-\mathrm{e}^{-s/t})^{n-1}\mathrm{d}s.
\end{aligned}$$

By making the replacement $u = 1 - \mathrm{e}^{-s/t}$ in the integrals, we obtain

$$\begin{aligned}
f_{\tau_{n-1}}(t) &= -n\int_0^1 (1-u)^{nt-1}u^{n-1}\mathrm{d}u - n^2 t\int_0^1 \ln(1-u)(1-u)^{nt-1}u^{n-1}\mathrm{d}u \\
&= -n\left[B(n,nt) + nt\frac{\partial B(n,nt)}{\partial(nt)}\right].
\end{aligned}$$

Since

$$\frac{\partial B(n,nt)}{\partial(nt)} = B(n,nt)[\psi(nt) - \psi(n+nt)],$$

where $\psi(z) = \frac{\mathrm{d}}{\mathrm{d}z}\ln z = \frac{\Gamma'(z)}{\Gamma(z)}$ is the digamma function, it follows that

$$f_{\tau_{n-1}}(t) = nB(n,nt)\{nt[\psi(n+nt) - \psi(nt)] - 1\}\,\mathbb{I}_{(0,\infty)}(t)\,. \tag{27}$$

Using the recurrence formula for the digamma function,

$$\psi(n+z) = \sum_{k=1}^{n}\frac{1}{n-k+z} + \psi(z)\,,\quad n \in \mathbb{N}\,,$$

(cf. Abramowitz and Stegun [18]), the probability density function of τ_{n-1}, given in (27), can be further simplified to

$$f_{\tau_{n-1}}(t) = nB(n,nt)\sum_{k=1}^{n-1}\frac{nt}{n-k+nt}\,\mathbb{I}_{(0,\infty)}(t)\,.$$

In Table A1, in Appendix A, critical values of τ_{n-1} are displayed. The critical values were obtained using Mathematica v12. In Figure 5, the probability density function of τ_{n-1} is plotted for some values of n.

For large values of n, an approximation can be used to obtain critical values of the statistic τ_{n-1}. In recalling that $Z_{n:n} - Z_{1:n} \overset{d}{=} Z_{n-1:n-1}$, the limit distribution of the sequence of maxima $(Z_{n-1:n-1} - \ln(n-1))_{n\geq 2}$ is a standard Gumbel distribution, and therefore,

$$Z_{n:n} - Z_{1:n} - \ln(n-1) \xrightarrow[n\to\infty]{d} V \sim \text{Gumbel}(0,1)\,.$$

Moreover, as $\lim\limits_{n\to\infty}\frac{\ln n}{\ln(n-1)} = 1$, in using the fact that $\frac{Z_{n-1:n-1}}{\ln n} \overset{P}{\to} 1$ and $n\,Z_{1:n} \overset{d}{=} Z$, it follows from Slutsky's theorem [19] that

$$n\ln n\,\tau_{n-1} \xrightarrow[n\to\infty]{d} Z \sim \text{Exponential}(0,1)\,.$$

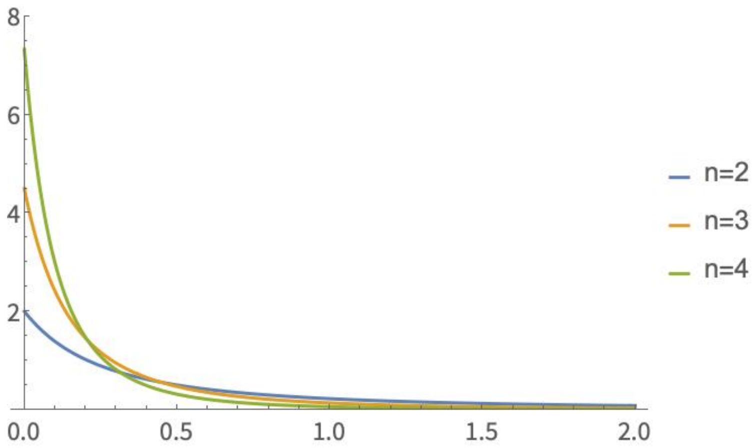

Figure 5. Probability density function of τ_{n-1}.

4.3. *Internal Studentization Using Sums of Spacings*

As the minimum $X_{1:n}$ and/or the maximum $X_{n:n}$ can be outliers, according to Hampel's [20] breakdown point concept, the statistic τ_{n-1}, defined in (26), can be unreliable. For an overview of robustness issues, see Rochetti [21], and in what concerns the duality of perspectives for extreme values/outliers, see Bhattacharya et al. [22]. For these reasons, it is more conservative to use instead the internally Studentized statistic

$$\tau_{n-1;i,k} = \frac{\overline{X}_n - \lambda}{X_{k:n} - X_{i:n}}\,, \tag{28}$$

with $1 < i < k < n$.

The sample mean $\overline{X}_n$ is a sufficient and complete estimator of λ, and the distribution of $\tau_{n-1;i,k}$ does not depend on the nuisance scale parameter δ. Hence, from Basu's [23] theorem, we conclude that $\overline{X}_n$ and $\tau_{n-1;i,k}$ are independent.

Noticing also that $\tau_{n-1;i,k} \stackrel{d}{=} \frac{\overline{Z}_n}{Z_{k:n}-Z_{i:n}}$, from the equality

$$Z_{k:n} - Z_{i:n} = \frac{\overline{Z}_n}{\tau_{n-1;i,k}}\,,$$

we must have $f_{Z_{k:n}-Z_{i:n}} = f_{\frac{\overline{Z}_n}{\tau_{n-1;i,k}}}$. Hence, from $\overline{Z}_n \sim \text{Gamma}\left(n, \frac{1}{n}\right)$ and $Z_{k:n} - Z_{i:n} \stackrel{d}{=} Z_{k-i:n-i}$, we have for $x > 0$,

$$\frac{\Gamma(n+1-i)}{\Gamma(k-i)\Gamma(n+1-k)} \mathrm{e}^{-x(n+1-k)} \left(1-\mathrm{e}^{-x}\right)^{k-i-1} = \int_0^\infty f_{\overline{Z}_n}(xt) f_{\tau_{n-1;i,k}}(t)\, t\, \mathrm{d}t\,,$$

i.e.,

$$\frac{\Gamma(n+1-i)}{\Gamma(k-i)\Gamma(n+1-k)} \mathrm{e}^{-x(n+1-k)} \left(1-\mathrm{e}^{-x}\right)^{k-i-1} = \int_0^\infty \frac{n(nxt)^{n-1}\mathrm{e}^{-nxt}}{\Gamma(n)} f_{\tau_{n-1;i,k}}(t)\, t\, \mathrm{d}t\,.$$

Thus, denoting $\mathcal{L}(f;x)$ the Laplace transform of function f at x, we obtain

$$\mathcal{L}(t^n f_{\tau_{n-1;i,k}}(t);x) = \frac{\Gamma(n)\Gamma(n+1-i)}{n\Gamma(k-i)\,\Gamma(n+1-k)} \frac{\mathrm{e}^{-x(n+1-k)/n}(1-\mathrm{e}^{-x/n})^{k-i-1}}{x^{n-1}}\,,$$

and in denoting $\mathcal{L}^{-1}(g;t)$ the inverse Laplace transform of function g at t, it follows that

$$f_{\tau_{n-1;i,k}}(t) = \frac{\Gamma(n)\Gamma(n+1-i)}{n\Gamma(k-i)\,\Gamma(n+1-k)}\, t^{-n} \mathcal{L}^{-1}\left(\frac{\mathrm{e}^{-x(n+1-k)/n}(1-\mathrm{e}^{-x/n})^{k-i-1}}{x^{n-1}};t\right). \tag{29}$$

For example, if $n = 2$, $i = 1$ and $k = 2$,

$$f_{\tau_{1;1,2}}(t) = \frac{1}{2t^2}\mathcal{L}^{-1}\left(\frac{e^{-x/2}}{x};t\right) = \frac{1}{2t^2}\,\mathbb{I}_{(\frac{1}{2},\infty)}(t)\,,$$

i.e., $\tau_{1;1,2} \sim \text{Pareto}(1, \frac{1}{2})$.

The most interesting scenario for (29) is when the integers n, i, and k satisfy the condition $n+1-k = k-i-1$, which occurs in the following cases:

- $n = 3j-1$, $i = j-1$ and $k = 2j$;
- $n = 3j$, $i = j$ and $k = 2j+1$;
- $n = 3j+1$, $i = j+1$ and $k = 2j+2$.

If this happens, then

$$f_{\tau_{n-1;i,k}}(t) = \frac{\Gamma(n)\Gamma(2j+1)}{n\Gamma(j)\Gamma(j+1)}\,\frac{1}{t^n}\,\mathcal{L}^{-1}\left(\frac{1}{x^{k-2}}\left(\frac{e^{-x/n}(1-e^{-x/n})}{x}\right)^j;t\right).$$

Since

$$\mathcal{L}^{-1}\left(\frac{1}{x^{k-2}};t\right) = \frac{t^{k-3}}{\Gamma(k-2)} \quad \text{and} \quad \mathcal{L}^{-1}\left(\frac{e^{-x/n}(1-e^{-x/n})}{x};t\right) = \mathbb{I}_{(\frac{1}{n},\frac{2}{n})}(t)\,,$$

from the fact that the inverse Laplace transform of a product is the convolution of the inverse Laplace transforms of the factors, we obtain, for example, for $n = 3$, $i = 1$, and $k = 3$,

$$f_{\tau_{2;1,3}}(t) = \begin{cases} 0 & ,\ t < \frac{1}{3} \\ \frac{4(t-\frac{1}{3})}{3t^3} & ,\ \frac{1}{3} \le t < \frac{2}{3} \\ \frac{4}{9t^3} & ,\ t \ge \frac{2}{3} \end{cases},$$

and for $n = 4$, $i = 2$, and $k = 4$,

$$f_{\tau_{3;2,4}}(t) = \begin{cases} 0 & ,\ t < \frac{1}{4} \\ \frac{3(4t-1)^2}{32t^4} & ,\ \frac{1}{4} \le t < \frac{1}{2} \\ \frac{3(8t-3)}{32t^4} & ,\ t \ge \frac{1}{2} \end{cases}.$$

In Table A2, in Appendix A, critical values of $\tau_{n-1;i,k}$ are supplied for integers n, i, and k satisfying $n+1-k = k-i-1$. The critical values of $\tau_{n-1;i,k}$ were also obtained with Mathematica v12. The probability density functions of $\tau_{2;1,3}$ and $\tau_{3;2,4}$ are shown in Figure 6.

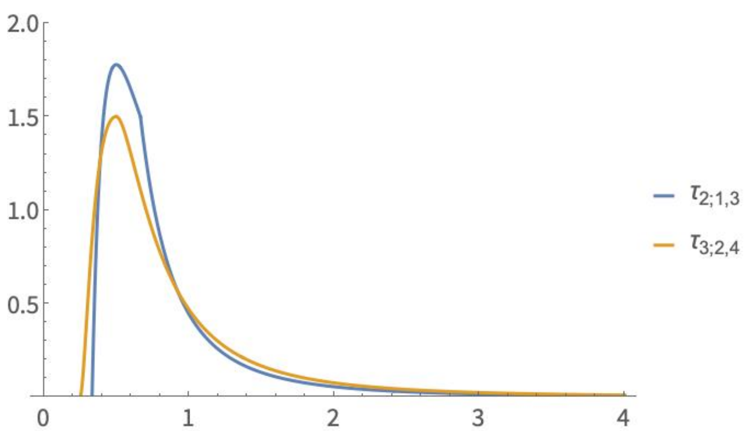

Figure 6. Probability density functions of $\tau_{2;1,3}$ and $\tau_{3;2,4}$.

4.4. Comparing the Locations of Two Exponential Populations with Equal Dispersions

For two exponential populations $X \sim \text{Exponential}(\lambda_1, \delta_1)$ and $Y \sim \text{Exponential}(\lambda_2, \delta_2)$, it may be of interest to make inferences on $\lambda_1 - \lambda_2$. For this purpose, Theorem 1 can be helpful.

Theorem 1. *Let X_1 and X_2 be two independent random variables such that $X_1 \sim \text{Exponential}(\lambda_1, \delta_1)$ and $X_2 \sim \text{Exponential}(\lambda_2, \delta_2)$. Then, $X_1 - X_2$ has an asymmetric Laplace distribution with probability density function*

$$f_{X_1 - X_2}(x) = \begin{cases} \frac{1}{\delta_1+\delta_2}\, \mathrm{e}^{\frac{x-(\lambda_1-\lambda_2)}{\delta_2}} & , \; x \le \lambda_1 - \lambda_2 \\ \frac{1}{\delta_1+\delta_2}\, \mathrm{e}^{-\frac{x-(\lambda_1-\lambda_2)}{\delta_1}} & , \; x > \lambda_1 - \lambda_2 \end{cases}.$$

Proof. Let $Y = (X_1 - \lambda_1) - (X_2 - \lambda_2)$. Since $X_1 - \lambda_1 \sim \text{Exponential}(0, \delta_1)$ and $X_2 - \lambda_2 \sim \text{Exponential}(0, \delta_2)$, for $y \le 0$,

$$f_Y(y) = \int_{-y}^{\infty} f_{X_1-\lambda_1}(y+x) f_{X_2-\lambda_2}(x)\,\mathrm{d}x = \frac{1}{\delta_1\delta_2} \int_{-y}^{\infty} \mathrm{e}^{-\frac{y+x}{\delta_1}}\,\mathrm{e}^{-\frac{x}{\delta_2}}\,\mathrm{d}x = \frac{1}{\delta_1+\delta_2}\,\mathrm{e}^{\frac{y}{\delta_2}},$$

and for $y > 0$,

$$f_Y(y) = \int_{0}^{\infty} f_{X_1-\lambda_1}(y+x) f_{X_2-\lambda_2}(x)\,\mathrm{d}x = \frac{1}{\delta_1\delta_2} \int_{0}^{\infty} \mathrm{e}^{-\frac{y+x}{\delta_1}}\,\mathrm{e}^{-\frac{x}{\delta_2}}\,\mathrm{d}x = \frac{1}{\delta_1+\delta_2}\,\mathrm{e}^{-\frac{y}{\delta_1}}.$$

Hence, from $X_1 - X_2 \overset{d}{=} Y + (\lambda_1 - \lambda_2)$, it follows that

$$f_{X_1 - X_2}(x) = f_Y(x - (\lambda_1 - \lambda_2)) = \begin{cases} \frac{1}{\delta_1+\delta_2}\, \mathrm{e}^{\frac{x-(\lambda_1-\lambda_2)}{\delta_2}} & , \; x \le \lambda_1 - \lambda_2 \\ \frac{1}{\delta_1+\delta_2}\, \mathrm{e}^{-\frac{x-(\lambda_1-\lambda_2)}{\delta_1}} & , \; x > \lambda_1 - \lambda_2 \end{cases}.$$

□

If in Theorem 1 it is considered that $\delta_1 = \delta_2 = \delta$, the symmetric Laplace distribution with location parameter $\lambda_1 - \lambda_2 \in \mathbb{R}$, and scale parameter $\delta > 0$ is obtained. For more details on asymmetric Laplace random variables, see Brilhante and Kotz [24].

When dealing with two populations with equal dispersions, inferences on the difference between the locations of the two populations can be carried out using an externally Studentized statistic as described below.

Let $\boldsymbol{X} = (X_1, \ldots, X_{n_1})$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_{n_2})$ be two independent random samples with parent distributions $X \sim \text{Exponential}(\lambda_1, \delta)$ and $Y \sim \text{Exponential}(\lambda_2, \delta)$, respectively. Considering that

$$\frac{X_{1:n_1} - \lambda_1}{\delta} \sim \text{Exponential}\big(0, \tfrac{1}{n_1}\big) \quad \text{and} \quad \frac{Y_{1:n_2} - \lambda_2}{\delta} \sim \text{Exponential}\big(0, \tfrac{1}{n_2}\big),$$

from Theorem 1,

$$U = \frac{X_{1:n_1} - Y_{1:n_2} - (\lambda_1 - \lambda_2)}{\delta}$$

has probability density function

$$f_U(x) = \begin{cases} \frac{n_1 n_2}{N}\, \mathrm{e}^{n_2 x} & , \; x \le 0 \\ \frac{n_1 n_2}{N}\, \mathrm{e}^{-n_1 x} & , \; x > 0 \end{cases},$$

with $N = n_1 + n_2$.

On the other hand, as

$$n_1\left(\overline{X}_{n_1}-X_{1:n_1}\right)=\sum_{k=2}^{n_1}(n_1+1-k)\left(X_{k:n_1}-X_{k-1:n_1}\right)\sim\text{Gamma}(n_1-1,\delta)\,,$$

and, similarly, $n_2\left(\overline{Y}_{n_2}-Y_{1:n_2}\right)\sim\text{Gamma}(n_2-1,\delta)$, we have

$$V=\frac{\widehat{\delta}}{\delta}=\frac{\left[n_1\left(\overline{X}_{n_1}-X_{1:n_1}\right)+n_2\left(\overline{Y}_{n_2}-Y_{1:n_2}\right)\right]/N}{\delta}\sim\text{Gamma}\left(N-2,\tfrac{1}{N}\right).$$

The independence of spacings in exponential populations ensures that U and V are independent, and therefore, the externally Studentized statistic $W_{n_1-1,n_2-1}=\frac{U}{V}$, i.e.,

$$W_{n_1-1,n_2-1}=\frac{X_{1:n_1}-Y_{1:n_2}-(\lambda_1-\lambda_2)}{\left[n_1(\overline{X}_{n_1}-X_{1:n_1})+n_2(\overline{Y}_{n_2}-Y_{1:n_2})\right]/N}\,,$$

can be used to make inferences on $\lambda_1-\lambda_2$.

As for the probability density function of W_{n_1-1,n_2-1}, for $w\le 0$,

$$\begin{aligned}f_{W_{n_1-1,n_2-1}}(w)&=\int_{-\infty}^{\infty}f_U(wx)f_V(x)|x|\mathrm{d}x\\&=\frac{n_1n_2N^{N-3}}{\Gamma(N-2)}\int_0^{\infty}x^{N-2}\mathrm{e}^{-(N-n_2w)x}\mathrm{d}x\\&=\frac{n_1n_2(N-2)}{N^2\left(1-\frac{n_2}{N}w\right)^{N-1}}\,,\end{aligned}$$

and for $w>0$,

$$\begin{aligned}f_{W_{n_1-1,n_2-1}}(w)&=\int_{-\infty}^{\infty}f_U(wx)f_V(x)|x|\mathrm{d}x\\&=\frac{n_1n_2N^{N-3}}{\Gamma(N-2)}\int_0^{\infty}x^{N-2}\mathrm{e}^{-(N+n_1w)x}\mathrm{d}x\\&=\frac{n_1n_2(N-2)}{N^2\left(1+\frac{n_1}{N}w\right)^{N-1}}\,.\end{aligned}$$

Therefore,

$$f_{W_{n_1-1,n_2-1}}(w)=\begin{cases}\frac{n_1n_2(N-2)}{N^2\left(1-\frac{n_2}{N}w\right)^{N-1}} & ,\ w\le 0\\ \frac{n_1n_2(N-2)}{N^2\left(1+\frac{n_1}{N}w\right)^{N-1}} & ,\ w>0\end{cases}.\tag{30}$$

From (30), $\mathbb{P}(W\le 0)=\frac{n_1}{N}$ and $\mathbb{P}(W>0)=\frac{n_2}{N}=1-\frac{n_1}{N}$. Notice that $\frac{n_1}{N}$ is the proportion of observations in the combined sample that comes from the population X.

The critical value $W_{n_1-1,n_2-1;\alpha}$, $0<\alpha<1$, of W_{n_1-1,n_2-1} is

$$W_{n_1-1,n_2-1;\alpha}=\begin{cases}\frac{N}{n_2}\left[1-\left(\frac{n_1}{N\alpha}\right)^{\frac{1}{N-2}}\right] & ,\ 0<\alpha\le\frac{n_1}{N}\\ \frac{N}{n_1}\left[\left(\frac{n_2}{N(1-\alpha)}\right)^{\frac{1}{N-2}}-1\right] & ,\ \frac{n_1}{N}<\alpha<1\end{cases}.$$

If a balanced design is considered instead, i.e., $n_1=n_2=n$ in (30), denoting, for simplicity, $W_{n-1,n-1}=W_{n-1}$, then

$$f_{W_{n-1}}(w)=\frac{n-1}{2\left(1+\frac{|w|}{2}\right)^{2n-1}}\,\mathbb{I}_{\mathbb{R}}(w).\tag{31}$$

The $(1-\alpha)$-th critical value of W_{n-1} is $W_{n-1;1-\alpha} = 2\left(\frac{1}{(2\alpha)^{1/2(n-1)}} - 1\right)$, $\alpha < 0.5$. (The distribution of W_{n-1} is symmetric around zero.)

In Figure 7, the probability density function of W_{n_1,n_2} is plotted for some values of n_1 and n_2.

Observe that if the two populations have unequal dispersions, i.e., $\delta_1 \neq \delta_2$, the distribution of W_{n_1-1,n_2-1} depends on the nuisance scale parameters δ_1 and δ_2, and so, it cannot be used to make inferences on $\lambda_1 - \lambda_2$, unless a Satterthwaite [15] approximation is considered (for more details, see Section 4.6).

4.5. Analysis of Spacings (ANOSp) for Testing Homogeneity of $k > 2$ Locations of Exponential Populations with Equal Dispersions

Let $\boldsymbol{X_1} = (X_{11}, \ldots, X_{1n_1}), \ldots, \boldsymbol{X_k} = (X_{k1}, \ldots, X_{kn_k})$ be $k > 2$ independent random samples, with $X_{ij} \sim \text{Exponential}(\lambda_i, \delta_i)$, $i = 1, \ldots, k$, $j = 1, \ldots, n_i$. For what follows, we shall use the notations $\overline{X}_i$ for the sample mean and $X^{(i)}_{j:n_i}$ for the j-th ascending order statistic, $j = 1, \ldots, n_i$, of the i-th random sample, $i = 1, \ldots, k$. We shall also use the notations $\overline{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}$ and $X_{1:N} = \min_{\substack{1\le i\le k \\ 1\le j\le n_i}} X_{ij} = \min_{1\le i\le k} X^{(i)}_{1:n_i}$, with $N = n_1 + \cdots + n_k$ denoting the size of the combined samples.

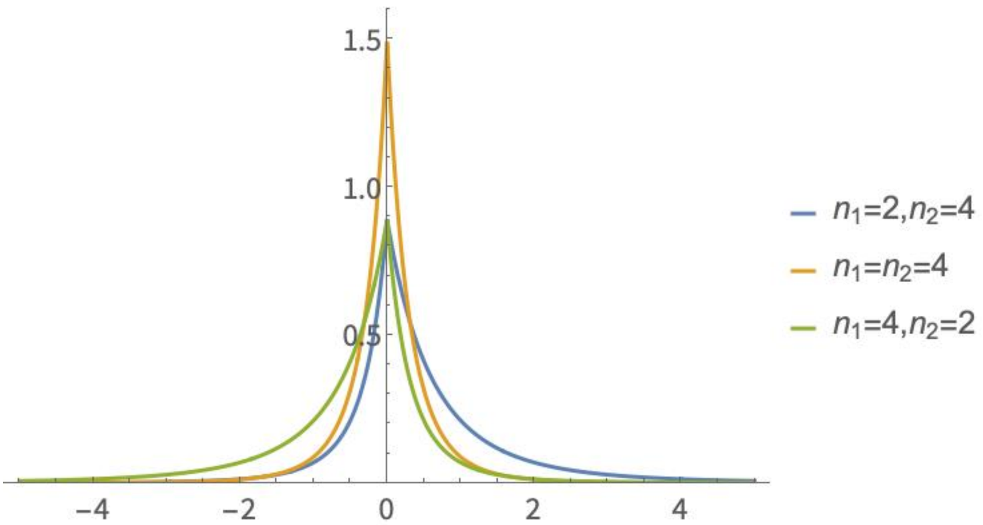

Figure 7. Probability density function of W_{n_1-1,n_2-1}.

The maximum likelihood estimators of the location and scale parameters of the individual populations are

$$\widehat{\lambda}_i = X^{(i)}_{1:n_i} \sim \text{Exponential}\left(\lambda_i, \tfrac{\delta_i}{n_i}\right) \quad \text{and} \quad \widehat{\delta}_i = \overline{X}_i - X^{(i)}_{1:n_i} \sim \text{Gamma}\left(n_i - 1, \tfrac{\delta_i}{n_i}\right),$$

$i = 1, \ldots, k$. Our interest lies in testing the homogeneity of locations of the populations, i.e., testing

$$H_0 : \lambda_1 = \cdots = \lambda_k \,(= \lambda) \quad \text{vs.} \quad H_A : \lambda_i \neq \lambda_j\,, \quad \text{for some } i, j \in \{1,, \ldots, k\}\,, i \neq j\,.$$

For the time being, we shall assume that the populations have equal dispersions, i.e., $\delta_1 = \ldots = \delta_k = \delta$.

In this setting, a parallelism with a one-way ANOVA scheme can be made, since the Total Sum of Spacings (TSSp) can be split into a Between Sum of Spacings (BSSp) and a Within Sum of Spacings (WSSp) as follows:

$$\begin{aligned}\text{TSSp} &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - X_{1:N}\right) \\ &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left[\left(X_{ij} - X_{1:n_i}^{(i)}\right) + \left(X_{1:n_i}^{(i)} - X_{1:N}\right)\right] \\ &= \sum_{i=1}^{k} n_i\left(X_{1:n_i}^{(i)} - X_{1:N}\right) + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - X_{1:n_i}^{(i)}\right) \\ &= \text{BSSp} + \text{WSSp}\,.\end{aligned}$$

In noticing that

$$\text{WSSp} = \sum_{i=1}^{k}\sum_{j=2}^{n_i}\left(X_{j:n_i}^{(i)} - X_{1:n_i}^{(i)}\right) = \sum_{i=1}^{k}\sum_{j=2}^{n_i}(n_i + 1 - j)\left(X_{j:n_i}^{(i)} - X_{j-1:n_i}^{(i)}\right) \sim \text{Gamma}(N-k, \delta)\,,$$

the statistic $\frac{\text{WSSp}}{N-k}$ is an unbiased estimator of δ, regardless of the validity of H_0.

On the other hand, under H_0, the random variables $Y_i = n_i X_{1:n_i}^{(i)} \sim \text{Exponential}(\lambda, \delta)$, $i = 1, \ldots, k$, are independent. Thus, from

$$\text{BSSp} = \sum_{i=1}^{k} n_i\left(X_{1:n}^{(i)} - X_{1:N}\right) = \sum_{i=1}^{k} n_i X_{1:n_i}^{(i)} - N X_{1:N}$$

and $N X_{1:N} \stackrel{d}{=} k Y_{1:k}$, with $Y_{1:k} = \min\{Y_1, \ldots, Y_k\}$, it follows that

$$\text{BSSp} \stackrel{d}{=} \sum_{i=1}^{k} Y_i - k Y_{1:k} = k\left(\overline{Y}_k - Y_{1:k}\right) \sim \text{Gamma}(k-1, \delta)\,.$$

Thus, under H_0, the statistic $\frac{\text{BSSp}}{k-1}$ is an unbiased estimator of δ.

From the above, and under H_0, the F-statistic

$$F = \frac{\text{MBSSp}}{\text{MWSSp}} = \frac{\frac{\text{BSSp}}{k-1}}{\frac{\text{WSSp}}{N-k}} = \frac{\frac{\text{BSSp}/(2\delta)}{2(k-1)}}{\frac{\text{WSSp}/(2\delta)}{2(N-k)}} \sim F_{2(k-1),2(N-k)} \tag{32}$$

can detect gross departures from the null hypothesis. Notice that the independence of spacings in exponential populations guarantees the independence of BSSp and WSSp.

An ANOSp table, similar to a one-way ANOVA table, can be shown in this context (see Table 1).

Table 1. ANOSp table.

Sum of Spacings	**df**	**Mean of Sum of Spacings**	***F*-Statistic**
$\text{BSSp} = \sum_{i=1}^{k} n_i\left(X_{1:n_i}^{(i)} - X_{1:N}\right)$	$2(k-1)$	$\text{MBSSp} = \frac{\text{BSSp}}{2(k-1)}$	$F = \frac{\text{MBSSp}}{\text{MWSSp}} = \frac{\frac{\text{BSSp}}{k-1}}{\frac{\text{WSSp}}{N-k}}$
$\text{WSSp} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij} - X_{1:n_i}^{(i)}\right)$	$2(N-k)$	$\text{MWSSp} = \frac{\text{WSSp}}{2(N-k)}$	

4.6. Analysis of Spacings (ANOSp) for Testing Homogeneity of $k > 2$ Locations of Exponential Populations with Unequal Dispersions

If the exponential populations have unequal dispersions, the statistic defined in (32) is useless, because its distribution depends now on the nuisance scale parameters of the individual populations. However, a Satterthwaite approximation can be considered to eliminate its distribution's dependence on those nuisance parameters.

Satterthwaite [15] showed that if $Y_1, \ldots, Y_n$ are independent random variables such that $Y_i \sim \chi^2_{\nu_i}$, $i = 1, \ldots, n$, the linear combination $\sum_{i=1}^n a_i Y_i$, with $a_i \in \mathbb{R}$, $i = 1, \ldots, n$, can be approximated using $\frac{Y}{\nu}$, where $Y \sim \chi^2_{\nu}$, i.e., $\sum_{i=1}^n a_i Y_i \approx \frac{Y}{\nu}$, where the degree of freedom ν is estimated using the estimator

$$\widetilde{\nu} = \frac{\left(\sum_{i=1}^n a_i Y_i\right)^2}{\sum_{i=1}^n \frac{a_i^2}{\nu_i} Y_i^2}\,. \tag{33}$$

Since under H_0, the distribution of MBSSp does not depend on λ, we shall assume, without loss of generality, that $\lambda = 0$. We shall also assume that from $X_{1:N} = \min\limits_{1 \le i \le k} X^{(i)}_{1:n_i}$, in particular, $X_{1:N} = X^{(1)}_{1:n_1}$, and therefore,

$$\begin{aligned}
\text{MBSSp} &= \frac{1}{k-1}\sum_{i=1}^k n_i\left(X^{(i)}_{1:n_i} - X_{1:N}\right)\\
&= \frac{1}{k-1}\sum_{i=1}^k n_i\left(X^{(i)}_{1:n_i} - X^{(1)}_{1:n_1}\right)\\
&= \frac{n_1 - N}{k-1} X^{(1)}_{1:n_1} + \sum_{i=2}^k \frac{n_i}{k-1} X^{(i)}_{1:n_i}\\
&= \frac{n_1 - N}{k-1}\widehat{\lambda}_1 + \sum_{i=2}^k \frac{n_i}{k-1}\widehat{\lambda}_i\,,
\end{aligned}$$

which can be expressed as

$$\text{MBSSp} = \frac{\delta_1(n_1 - N)}{2n_1(k-1)} Y_1 + \sum_{i=2}^k \frac{\delta_i}{2(k-1)}\, Y_i \approx \frac{U}{\nu_1}\,,$$

with $Y_i = \frac{2n_i}{\delta_i}\widehat{\lambda}_i \sim \chi^2_2$, $i = 1, \ldots, k$, and $U \sim \chi^2_{\nu_1}$. In using $a_1 = \frac{\delta_1(n_1-N)}{2n_1(k-1)}$ and $a_i = \frac{\delta_i}{2(k-1)}$, $i = 2, \ldots, k$, in Formula (33), the parameter ν_1 is estimated by

$$\widetilde{\nu}_1 = \frac{2\left(\sum_{i=1}^k n_i\,\widehat{\lambda}_i - N\,\widehat{\lambda}_1\right)}{(n_1 - N)^2\,\widehat{\lambda}_1^2 + \sum_{i=2}^k n_i^2\,\widehat{\lambda}_i^2}\,. \tag{34}$$

On the other hand, noticing that

$$\begin{aligned}
\text{MWSSp} &= \frac{1}{N-k}\sum_{i=1}^k \sum_{j=1}^{n_i}\left(X_{ij} - X^{(i)}_{1:n_i}\right)\\
&= \frac{1}{N-k}\sum_{i=1}^k n_i\left(\overline{X}_i - X^{(i)}_{1:n_i}\right)\\
&= \sum_{i=1}^k \frac{n_i}{N-k}\widehat{\delta}_i\,,
\end{aligned}$$

with $\widehat{\delta}_i \sim \text{Gamma}(n_i - 1, \delta_i)$, $i = 1, \ldots, k$, and then considering $Y_i = \frac{2}{\delta_i}\widehat{\delta}_i \sim \chi^2_{2(n_i-1)}$, $i = 1, \ldots, k$, we have

$$\text{MWSSp} = \sum_{i=1}^k \frac{n_i \delta_i}{2(N-K)} Y_i \approx \frac{V}{\nu_2}\,,$$

with $V \sim \chi^2_{\nu_2}$. In using $a_i = \frac{n_i\delta_i}{2(N-K)}$, $i = 1, \ldots, n$, in Formula (33), the parameter ν_2 is estimated by

$$\widetilde{\nu}_2 = \frac{2\left(\sum_{i=1}^k n_i\,\widehat{\delta}_i\right)^2}{\sum_{i=1}^k \frac{n_i^2}{n_i - 1}\,\widehat{\delta}_i^2}\,. \tag{35}$$

Consequently, under H_0, U and V are independent, and therefore,

$$F = \frac{\text{MBSSp}}{\text{MWSSp}} \approx \frac{U/\nu_1}{V/\nu_2} \sim F_{\nu_1,\nu_2},$$

where the degrees of freedom ν_1 and ν_2 of the F-statistic are estimated using (34) and (35), respectively.

5. Conclusions

The exact distribution of externally Studentized statistics in Gaussian samples and in exponential samples are readily obtained. The members of the location-scale exponential family with a pair of sufficient statistics for the parameters are as follows:

1. $X \sim \text{Gaussian}(\mu, \sigma)$, when the support is the real line;
2. $Y \sim \text{Exponential}(\lambda, \delta)$, when the support is the half-line (λ, ∞) (or the reverted exponential when the support is $(-\infty, \lambda)$, in which case the pair of sufficient statistics is $\left(Y_{n:n}, \sum_{k=1}^{n} Y_k\right)$, and the maximum likelihood estimators of the parameters are $\widehat{\lambda} = Y_{n:n}$ and $\widehat{\delta} = Y_{n:n} - \overline{Y}_n$);
3. $W \sim \text{Uniform}(\lambda, \lambda + \delta)$ when the support is a segment, in which case the pair of sufficient statistics is$(W_{1:n}, W_{n:n})$ and the maximum likelihood estimators of the parameters are $\widehat{\lambda} = W_{1:n}$ and $\widehat{\delta} = W_{n:n} - W_{1:n}$.

While for Gaussian samples and for exponential samples, the probability density functions of externally Studentized statistics are easily obtained, in the case of uniform samples, this is not possible. However, as uniform random variables are symmetric by default, the approximation under the smoothness hypothesis 1 holds, although it is known from Hendriks [11] that this is the worst framework to consider. Alternatively, as uniform samples can be logarithmically transformed into exponential data, inferences on the location parameter(s) can be dealt with using the results of Section 4.

On the other hand, it is, in general, impossible to obtain exact results for internally Studentized statistics. As Hotelling [25], Efron [10], and Lehman [26] have shown, symmetry for the parent distribution is a useful property for Studentizations. The investigation of the approximation resulting from the smoothness hypothesis (17) for symmetric parents, inspired by the Gaussian case, shows that even in the worst $\text{Uniform}(-1, 1)$ case, which was identified by Hendriks et al. [11] as such, the approximation in the tails is quite good, and therefore, it is usable for inferences on the location parameter.

With regard to asymmetric exponential parent distributions, from the independence of spacings, exact results either for one sample inferences on the location parameter or for the comparison of $k \geq 2$ location parameters assuming equal dispersions are straightforward. An approximate solution for the unequal dispersions case, inspired by Satterthwaite's [15] treatment for comparing means in heteroscedastic Gaussian parents, has been presented.

Author Contributions: Conceptualization, M.F.B., D.P. and M.L.R.; methodology, M.F.B. and D.P.; investigation, M.F.B., D.P. and M.L.R.; writing—original draft preparation, M.F.B. and D.P.; writing—review and editing, M.F.B., D.P., and M.L.R.; supervision, M.F.B. and D.P.; project administration, M.F.B. and D.P.; funding acquisition, M.F.B., D.P. and M.L.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially financed by national funds through FCT (Fundação para a Ciência e a Tecnologia), Portugal, under project UIDB/00006/2020 (https://doi.org/10.54499/UIDB/00006/2020), and research grant UIDB/00685/2020 of the CEEAplA (Centre of Applied Economics Studies of the Atlantic), School of Business and Economics of the University of the Azores.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Tables with Critical Values

Table A1. Critical values of $\tau_{n-1} = \frac{X_{1:n}-\lambda}{X_{n:n}-X_{1:n}}$.

n	**0.001**	**0.005**	**0.01**	**0.025**	**0.05**	**0.1**	**0.9**	**0.95**	**0.975**	**0.99**	**0.995**	**0.999**
2	0.000501	0.002513	0.005051	0.012821	0.026316	0.055556	4.5	9.5	19.5	49.5	99.5	499.5
3	0.000222	0.001115	0.002240	0.005666	0.011562	0.024110	1	1.614763	2.486079	4.216991	6.168750	14.408050
4	0.000136	0.000684	0.001373	0.003470	0.007068	0.014677	0.5	0.75	1.067030	1.618460	2.164490	4.047390
5	0.000096	0.000482	0.000966	0.002441	0.004966	0.010291	0.319126	0.462948	0.635782	0.917745	1.179750	1.999380
6	0.000073	0.000366	0.000735	0.001855	0.003771	0.007805	0.228990	0.325932	0.438821	0.616237	0.775077	1.244660
7	0.000058	0.000292	0.000587	0.001481	0.003010	0.006224	0.176015	0.247467	0.328973	0.453985	0.563222	0.874548
8	0.000048	0.000242	0.000485	0.001224	0.002487	0.005140	0.141563	0.197318	0.259996	0.354495	0.435671	0.661185
9	0.000041	0.000205	0.000411	0.001038	0.002108	0.004354	0.117566	0.162821	0.213143	0.288052	0.351592	0.524816
10	0.000035	0.000177	0.000356	0.000897	0.001822	0.003762	0.1	0.137805	0.179488	0.240930	0.292540	0.431222
11	0.000031	0.000156	0.000312	0.000788	0.001599	0.003301	0.086647	0.118930	0.154283	0.205986	0.249079	0.363559
12	0.000028	0.000138	0.000278	0.000700	0.001422	0.002934	0.076193	0.104239	0.134782	0.179165	0.215923	0.312669
13	0.000025	0.000124	0.000249	0.000629	0.001277	0.002634	0.067811	0.092518	0.119300	0.158009	0.189898	0.273188
14	0.000022	0.000112	0.002292	0.000570	0.001157	0.002386	0.060958	0.082974	0.106745	0.140945	0.168995	0.241782
15	0.000021	0.000103	0.000206	0.000520	0.001056	0.002177	0.055261	0.075069	0.096381	0.126926	0.151881	0.216280
16	0.000019	0.000094	0.000189	0.000478	0.000970	0.001999	0.050459	0.068426	0.087699	0.115228	0.137644	0.195213
17	0.000017	0.000087	0.000175	0.000441	0.000896	0.001847	0.046363	0.062774	0.080332	0.105336	0.125636	0.177552
18	0.000016	0.000081	0.000162	0.0004010	0.000831	0.001714	0.042832	0.057913	0.074011	0.096874	0.115388	0.162561
19	0.000015	0.000076	0.000151	0.000382	0.000775	0.001597	0.039761	0.053694	0.068535	0.089564	0.106553	0.149697
20	0.000014	0.000071	0.000142	0.000357	0.000725	0.001495	0.037067	0.05000	0.063751	0.083191	0.098865	0.138550
21	0.000013	0.000066	0.000133	0.000336	0.000681	0.001404	0.034687	0.046743	0.059538	0.077593	0.092122	0.128812
22	0.000012	0.000063	0.000125	0.000316	0.000642	0.001322	0.032571	0.043851	0.055804	0.072641	0.086166	0.120238
23	0.000012	0.000059	0.000118	0.000299	0.000606	0.001249	0.030679	0.041268	0.052474	0.068232	0.080871	0.112640
24	0.000011	0.000056	0.000112	0.000283	0.000574	0.001183	0.028978	0.038950	0.049488	0.064285	0.076137	0.105864
25	0.000011	0.000053	0.000107	0.000269	0.000545	0.001123	0.027441	0.036857	0.046796	0.060733	0.071880	0.099790
30	0.000008	0.000042	0.000085	0.000213	0.000433	0.000891	0.021568	0.028883	0.036567	0.047281	0.055805	0.076986
50	0.000004	0.000022	0.000045	0.000113	0.000230	0.000472	0.011186	0.014887	0.018732	0.024031	0.028200	0.038399

Table A2. Critical values of $\tau_{n-1;i,k} = \frac{\overline{X}_n-\lambda}{X_{k:n}-X_{i:n}}$.

n	*i*	*k*	**0.001**	**0.005**	**0.01**	**0.025**	**0.05**	**0.1**	**0.9**	**0.95**	**0.975**	**0.99**	**0.995**	**0.999**
3	1	3	0.340957	0.350877	0.358697	0.375292	0.395936	0.429336	1.49071	2.10819	2.98142	4.71405	6.66667	14.9071
4	2	4	0.271553	0.289258	0.301567	0.325555	0.353308	0.395822	1.79672	2.60383	3.74135	5.99467	8.53244	19.2387
5	1	4	0.451275	0.481872	0.501286	0.536241	0.573179	0.624337	1.79163	2.31281	2.96466	4.08859	5.19617	9.00251
6	2	5	0.404304	0.439866	0.461814	0.500623	0.541023	0.597081	1.93635	2.52828	3.26727	4.54005	5.79359	10.0999
7	3	6	0.373282	0.412145	0.435759	0.477117	0.520069	0.580165	2.03545	2.67576	3.47458	4.84979	6.20392	10.855
8	2	6	0.495362	0.540184	0.566193	0.610203	0.654555	0.715025	1.93923	2.40243	2.94501	3.81446	4.6136	7.09137
9	3	7	0.469472	0.516524	0.543743	0.589912	0.636674	0.700813	2.00248	2.49299	3.0672	3.98691	4.83202	7.4518
10	4	8	0.449604	0.498404	0.526618	0.574574	0.623311	0.690388	2.05194	2.56382	3.1628	4.12196	5.00318	7.73456
11	3	8	0.540956	0.591963	0.620843	0.669181	0.717526	0.782997	1.96377	2.36465	2.81547	3.50589	4.11369	5.88409
12	4	9	0.523183	0.575748	0.605564	0.655587	0.70575	0.773818	1.99983	2.41515	2.88203	3.59686	4.22604	6.05843
13	5	10	0.508582	0.562481	0.593100	0.644558	0.696248	0.766474	2.02986	2.4572	2.93749	3.67270	4.31974	6.20397
14	4	10	0.580715	0.63532	0.665943	0.716882	0.767472	0.835382	1.95716	2.30928	2.6942	3.26553	3.75377	5.11685
15	5	11	0.567344	0.623225	0.654616	0.706903	0.758895	0.828738	1.98074	2.34184	2.73646	3.32208	3.82248	5.21933
16	6	12	0.555921	0.612927	0.644989	0.698448	0.751649	0.823149	2.00113	2.36998	2.7730	3.37101	3.88193	5.30806
17	5	12	0.61511	0.672086	0.703832	0.756351	0.808153	0.877114	1.94147	2.25603	2.59269	3.08098	3.48921	4.59421
18	6	13	0.604516	0.662577	0.694965	0.748586	0.801505	0.871973	1.95825	2.27896	2.62214	3.11981	3.53583	4.66184
19	7	14	0.595226	0.654259	0.687218	0.741813	0.795718	0.86751	1.97312	2.29927	2.64823	3.15422	3.57717	4.72183
20	6	14	0.645159	0.703758	0.736231	0.789683	0.842074	0.911308	1.92354	2.2086	2.50871	2.93625	3.28769	4.21667
21	7	15	0.636466	0.696005	0.729022	0.783394	0.836703	0.907156	1.93617	2.22574	2.53055	2.96473	3.32158	4.26483
22	8	16	0.628699	0.689089	0.722597	0.777795	0.831926	0.903471	1.94757	2.24118	2.555023	2.9904	3.35216	4.3083
23	7	16	0.671706	0.731426	0.764359	0.81832	0.870914	0.939971	1.90572	2.16709	2.43863	2.81995	3.12915	3.93126
24	8	17	0.664389	0.724931	0.758334	0.813078	0.866443	0.936514	1.91563	2.18046	2.45556	2.84183	3.15503	3.96745
25	9	18	0.657759	0.719053	0.752883	0.80834	0.862405	0.933397	1.92469	2.19266	2.47102	2.86182	3.17867	4.00052
30	10	21	0.711262	0.772902	0.806614	0.861427	0.914346	0.98310	1.87984	2.10779	2.33984	2.65864	2.91187	3.55044
50	16	34	—	0.885094	0.917836	0.970073	1.01943	1.082080	1.79297	1.95205	2.10773	2.31277	2.46933	2.84359

References

1. Student. The probable error of a mean. *Biometrika* **1908**, *6*, 1–25. [CrossRef]
2. Geary, R.C. Distribution of Student's ratio for nonnormal samples. *Suppl. J. R. Stat. Soc.* **1936**, *3*, 178–184. [CrossRef]
3. Darmois, G. Analyse générale des liaisons stochastiques: Étude particuière de l'analyse factorielle lineéaire. *Rev. L'Institut Int. Stat./Rev. Int. Stat. Inst.* **1953**, *21*, 2–8. [CrossRef]
4. Skitovich, V.P. Linear forms of independent random variables and the normal distribution law. *Izv. Akad. Nauk SSSR Ser. Mat.* **1954**, *18*, 185–200. (English Translation *Sel. Transl. Math. Stat. Probab.* **1962**, *2*, 211–218).
5. Koopman, B.O. On distributions admitting a sufficient statistic. *Trans. Am. Math. Soc.* **1936**, *39*, 399–409. [CrossRef]
6. Darmois, G. Sur les lois de probabilités à estimation exhaustive. *Comptes Rendus L'Académie Des Sci. Paris* **1935**, *200*, 1265–1266.
7. Pitman, E. Sufficient statistics and intrinsic accuracy. *Math. Proc. Camb. Philos. Soc.* **1936**, *32*, 567–579. [CrossRef]
8. Logan, B.F.; Mallows, C.L.; Rice, S.O.; Shepp, L.A. Limit distributions of self-normalized sums. *Ann. Probab.* **1973**, *1*, 788–809. [CrossRef]
9. Peña, V.H.D.L.; Lai, T.L.; Shao, Q.M. *Self-Normalized Processes. Limit Theory and Statistical Applications*; Springer: Berlin/Heidelberg, Germany, 2009.
10. Efron, B. Student's *t*-test under symmetry conditions. *J. Am. Stat. Assoc.* **1969**, *64*, 1278–1302. [CrossRef]
11. Hendriks, H.W.M.; Ijzerman-Boon, P.C.; Klaassen, C.A.J. Student's *t*-statistic under unimodal densities. *Austrian J. Stat.* **2006**, *35*, 131–141. [CrossRef]
12. Perlo, V. On the distribution of Student's ratio for samples of three drawn from the rectangular distribution. *Biometrika* **1933**, *25*, 203–204. [CrossRef]
13. Pexider, J.V. Notiz über Funkcionaltheoreme. *Monatsh. Math. Phys.* **1903**, *14*, 293–301. [CrossRef]
14. David, H.A. *Order Statistics*; Wiley: New York, NY, USA, 1981.
15. Satterthwaite, F.E. An approximate distribution of estimates of variance components. *Biom. Bull.* **1946**, *2*, 110–114. [CrossRef]
16. van Zwet, W. *Convex Transformations of Random Variables*, 7th ed.; Mathematical Centre: Amsterdam, The Netherlands, 1964.
17. van Zwet, W. Convex transformations: A new approach to skewness and kurtosis. *Statist. Neerl.* **1964**, *18*, 433–441. [CrossRef]
18. Abramowitz, M.; Stegun, I.A. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, 8th ed.; Dover: New York, NY, USA, 1972.
19. Slutsky, E. Über stochastische Asymptoten und Grenzwerte. *Metron* **1964**, *5*, 3–89. [CrossRef]
20. Hampel, F.R. Contributions to the Theory of Robust Estimation. Ph.D. Thesis, University of California, Berkeley, CA, USA, 1986.
21. Ronchetti, E. The main contributions of robust statistics to statistical science and a new challenge. *Metron* **2021**, *79*, 127–135. [CrossRef]
22. Bhattacharya, S.; Kamper, F.; Beirlant, J. Outlier detection based on extreme value theory and applications. *Scand. J. Stat.* **2024**, *50*, 1466–1502. [CrossRef]
23. Basu, D. On statistics independent of a complete sufficient statistic. *Sankhyā Indian J. Stat.* **1955**, *15*, 377–380. [CrossRef]
24. Brilhante, M.F.; Kotz, S. Infinite divisibility of the spacings of a Kotz–Kozubowski–Podgórski generalized Laplace model. *Stat. Probab. Lett.* **2008**, *78*, 2433–2436. [CrossRef]
25. Hotelling, H. The behavior of some standard statistical tests under nonstandard conditions. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability (Volume 1: Contributions to the Theory of Statistics)*; Neyman, J., Ed.; University of California Press: Berkeley, CA, USA, 1961; pp. 319–359.
26. Lehmann, E.L. "Student" and small-sample theory. *Stat. Sci.* **1999**, *14*, 418–426. [CrossRef]