Zeitschrift: Elemente der Mathematik

Herausgeber: Schweizerische Mathematische Gesellschaft

Band: 68 (2013)

Artikel: Two statistical coverage problems in estimating the variance of a population

Autor: Perstman, Wiebe R.

DOI: https://doi.org/10.5169/seals-515905

Elem. Math. 68 (2013) 154 – 166
0013-6018/13/040154-13
DOI 10.4171/EM/234

Elemente der Mathematik

Two statistical coverage problems in estimating the variance of a population

Wiebe R. Pestman

Wiebe R. Pestman obtained his doctor's degree in mathematics at the University of Groningen. Currently he works as a visiting professor at the mathematics department of the Federal University of Santa Catarina (Brasil). His interests are in functional analysis, harmonic analysis, probability and statistics.

1 Formulating the problems and setting the notation

Suppose a researcher draws a sample $X_1, X_2, \ldots, X_m$ from some population, and computes the corresponding variance in it. This in order to estimate the variance of the population from which the sample was drawn. Assume that the population in question has a Gaußian probability distribution. A second researcher draws, independently, a sample $Y_1, Y_2, \ldots, Y_n$ from the same population. He computes not only the corresponding variance in the sample, but also surrounds it by margins such as to get a 95% confidence interval for the population variance. Then what is the probability that this 95% confidence interval, generated by $Y_1, Y_2, \ldots, Y_n$, will cover the sample variance of $X_1, X_2, \ldots, X_m$? Below this probability will be denoted by P_n^m. As a second coverage problem, what is the probability that the 95% confidence interval generated by the $X_1, X_2, \ldots, X_m$ and that by the $Y_1, Y_2, \ldots, Y_n$ are disjoint? Below this probability will be denoted by Q_n^m. The aim

Zieht man zwei unabhängige Stichproben aus einer Grundgesamtheit mit Varianz σ^2, so kann man sich fragen, mit welcher Wahrscheinlichkeit das aus der zweiten Stichprobe gewonnene Vertrauensintervall für σ^2 den Wert der empirischen Varianz der ersten Stichprobe enthält. In ähnlicher Weise lässt sich fragen, wie gross die Wahrscheinlichkeit ist, dass die Vertrauensintervalle der beiden Stichproben für die Varianz disjunkt sind. In der vorliegenden Arbeit werden diese Fragen beantwortet für den Fall einer normalverteilten Grundgesamtheit. Insbesondere umfasst die Antwort jeden Stichprobenumfang und die entsprechende Asymptotik. Die Resultate werden angewandt auf statistische Tests, bei welchen die Überlappung der Vertrauensintervalle als Entscheidungskriterium dient.

of this paper is to get expressions for the values that can be taken on by P_n^m and Q_n^m and to get insight in their asymptotic behaviour. Of course one could also think about similar coverage problems when estimating the mean of a population rather than its variance. This has been studied in some detail in [9]. In the following, as a kind of a surprise, it will turn out that, asymptotically, the coverage probabilities P_n^n and Q_n^n are the same as their counterparts when estimating the mean.

2 Estimating the variance of a population

Let $X_1, X_2, \ldots, X_m$ be a sample from a population with variance σ^2. This variance σ^2 is then usually estimated through the so-called *sample variance* S^2, which is defined as

$$S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_m - \bar{X})^2}{m - 1}.$$

In the above the expression $\bar{X}$ stands for the sample mean, that is to say

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_m}{m}.$$

When sampling from a Gaußian population the random variables $\bar{X}$ and S^2 are statistically independent (see for example [7], [8]). By exploiting this result it can be proved that $(m-1)S^2/\sigma^2$ has a so-called χ^2-distribution with $m-1$ degrees of freedom. Generally a χ^2-distribution with n degrees of freedom is defined as being the probability distribution of a random variable of type

$$Z_1^2 + Z_2^2 + \cdots + Z_n^2$$

where the $Z_1, Z_2, \ldots, Z_n$ are independent random variables having a Gaußian distribution with mean 0 and variance 1. The fact that $(m-1)S^2/\sigma^2$ has a χ^2-distribution with $m-1$ degrees of freedom may be exploited to construct interval estimates at a prescribed coverage γ. To be more explicit in this, denote the quantile function of a χ^2-distribution with m degrees of freedom by q_m. Then, when using intervals with endpoints

$$\frac{(m-1)\,S^2}{q_{m-1}\big[\frac{1}{2}(1+\gamma)\big]} \quad \text{and} \quad \frac{(m-1)\,S^2}{q_{m-1}\big[\frac{1}{2}(1-\gamma)\big]}$$

the probability that they will cover the variance σ^2 of the population is precisely γ. Note that the number γ is in this context often referred to as the *confidence level* of the interval estimate. See for example [7] or [16] for more details in all this.

When drawing two samples $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ one may compute their variances S_X^2 and S_Y^2 and compare them by computing their quotient S_X^2/S_Y^2. If the samples are drawn from the same Gaußian population then this quotient S_X^2/S_Y^2 has a so-called F-distribution with $m-1$ and $n-1$ degrees of freedom in the numerator and denominator respectively. Generally an F-distribution with m and n degrees of freedom in the numerator and denominator is defined as being the probability distribution of a random variable of type

$$\frac{U/m}{V/n}$$

where U and V are independent variables having a χ^2-distribution with m and n degrees of freedom respectively. In the following such an F-distribution will be briefly referred to as an F_n^m-distribution. Its cumulative distribution function will be denoted as F_n^m. The families of F-distributions and χ^2-distributions play an important role in mathematical statistics (see for example [7]). In the following sections they will also play the central part in capturing the coverage probabilities P_n^m and Q_n^m in explicit expressions.

As a last notational convention, in the sections that follow the Greek capital Φ will stand for the cumulative distribution function of a standard Gaußian distribution, that is to say, a Gaußian distribution with mean 0 and variance 1. The quantile function of this distribution, that is the inverse of Φ, will be denoted by q.

3 A solution to the first coverage problem

Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ be independent samples from the same Gaußian population. As in the previous section, denote their corresponding sample variances by S_X^2 and S_Y^2 respectively. Then the endpoints of a confidence interval with coverage γ, generated by the sample $Y_1, Y_2, \ldots, Y_n$, are

$$\frac{(n-1)\,S_Y^2}{q_{n-1}\big[\frac{1}{2}(1+\gamma)\big]} \quad \text{and} \quad \frac{(n-1)\,S_Y^2}{q_{n-1}\big[\frac{1}{2}(1-\gamma)\big]}. \tag{1}$$

This interval will fail to cover the sample variance S_X^2 if either

$$S_X^2 < \frac{(n-1)\,S_Y^2}{q_{n-1}\big[\frac{1}{2}(1+\gamma)\big]} \quad \text{or} \quad \frac{(n-1)\,S_Y^2}{q_{n-1}\big[\frac{1}{2}(1-\gamma)\big]} < S_X^2.$$

These two events exclude each other. Exploiting the fact that S_X^2/S_Y^2 has an F_{n-1}^{m-1}-distribution, it is straightforward to derive that the probability that the interval given by (1) does not cover S_X^2 is given by

$$1 - P_n^m(\gamma) = F_{n-1}^{m-1}\left(\frac{n-1}{q_{n-1}\big[\frac{1}{2}(1+\gamma)\big]}\right) + 1 - F_{n-1}^{m-1}\left(\frac{n-1}{q_{n-1}\big[\frac{1}{2}(1-\gamma)\big]}\right).$$

It follows that

$$P_{n+1}^{m+1}(\gamma) = F_n^m\left(\frac{n}{q_n\big[\frac{1}{2}(1-\gamma)\big]}\right) - F_n^m\left(\frac{n}{q_n\big[\frac{1}{2}(1+\gamma)\big]}\right). \tag{2}$$

Table 1 below shows the probabilities P_n^m for a couple of values for m and n in a scenario where the coverage γ is set to 0.95. The probabilities are presented as percentages.

As to the asymptotics of the $P_n^m(\gamma)$, one has

$$\lim_{m\to\infty} P_n^m(\gamma) = \gamma \quad \text{for all } n \qquad \text{and} \qquad \lim_{n\to\infty} P_n^m(\gamma) = 0 \quad \text{for all } m.$$

These limits allow for easy intuitive explanations. The value of the left limit may be perceived as follows. With increasing m the S_X^2 converge (strongly) to σ_X^2. So the probability

m →	3	5	10	20	50	100	500	∞
$n = 3$	76.2	85.2	90.7	93.1	94.3	94.7	94.9	95
$n = 5$	68.1	79.6	87.8	91.7	93.8	94.4	94.9	95
$n = 10$	55.5	69.3	81.6	88.4	92.5	93.8	94.8	95
$n = 20$	42.4	56.6	72.0	82.6	90.0	92.6	94.5	95
$n = 50$	28.1	39.5	55.1	69.4	83.1	88.9	93.8	95
$n = 100$	20.1	29.0	42.2	56.4	73.8	83.3	92.6	95
$n = 500$	9.0	13.3	20.2	29.0	44.1	57.4	83.4	95
$n = \infty$	0	0	0	0	0	0	0	*

Table 1

that the confidence interval, at coverage γ, generated by the $Y_1, Y_2, \ldots, Y_n$, will cover S_X^2 may be expected to converge to the probability that it will cover σ_X^2. The latter probability, however, is γ by construction. As to the limit on the right, note that with increasing n the confidence intervals generated by the $Y_1, Y_2, \ldots, Y_n$ shrink to the singleton $\{\sigma_Y^2\}$. So the probability that these intervals will cover S_X^2 may be expected to converge to the probability that the singleton $\{\sigma_Y^2\}$ will cover S_X^2. The latter is precisely the probability that $S_X^2 = \sigma_Y^2$, which is 0 because S_X^2 has a continuous distribution.

Besides these two limits there is a limit of the P_n^m when walking along the diagonal of the table given above. To be more precise, it will turn out that the limit

$$\lim_{n\to\infty} P_n^n(\gamma)$$

exits and that it is equal to the limit of the P_n in [9], where corresponding coverage problems were studied in the estimation of a population mean. In §6 this result will be proved through analytic derivation.

4 A solution to the second coverage problem

In this section, for reasons that will become apparent in the last section, the second coverage problem will be solved in a slightly more general setting than proposed earlier. Namely, it will be assumed that the samples $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ are drawn from Gaußian populations with variances σ_X^2 and σ_Y^2 respectively. By assuming this the endpoints of the confidence interval for the population variance generated by the $Y_1, Y_2, \ldots, Y_n$ are

$$\frac{(n-1)\, S_Y^2}{q_{n-1}\left[\frac{1}{2}(1+\gamma)\right]} \quad \text{and} \quad \frac{(n-1)\, S_Y^2}{q_{n-1}\left[\frac{1}{2}(1-\gamma)\right]}.$$

Similarly, the endpoints of the interval generated by the $X_1, X_2, \ldots, X_m$ are

$$\frac{(m-1)\, S_X^2}{q_{m-1}\left[\frac{1}{2}(1+\gamma)\right]} \quad \text{and} \quad \frac{(m-1)\, S_X^2}{q_{m-1}\left[\frac{1}{2}(1-\gamma)\right]}.$$

The two intervals are disjoint if either

$$\frac{(m-1)\,S_X^2}{q_{m-1}\big[\frac{1}{2}(1-\gamma)\big]} < \frac{(n-1)\,S_Y^2}{q_{n-1}\big[\frac{1}{2}(1+\gamma)\big]}$$

or

$$\frac{(n-1)\,S_Y^2}{q_{n-1}\big[\frac{1}{2}(1-\gamma)\big]} < \frac{(m-1)\,S_X^2}{q_{m-1}\big[\frac{1}{2}(1+\gamma)\big]}\,.$$

It follows that, denoting the quotient σ_Y/σ_X by ρ, the probability $Q_{n+1}^{m+1}(\gamma,\rho)$ is given by

$$Q_{n+1}^{m+1}(\gamma,\rho) = F_n^m\left(\rho^2\frac{n}{m}\frac{q_m\big[\frac{1}{2}(1-\gamma)\big]}{q_n\big[\frac{1}{2}(1+\gamma)\big]}\right) + F_m^n\left(\frac{1}{\rho^2}\frac{m}{n}\frac{q_n\big[\frac{1}{2}(1-\gamma)\big]}{q_m\big[\frac{1}{2}(1+\gamma)\big]}\right). \qquad (3)$$

Table 2 below shows the probabilities Q_n^m for a couple of values for m and n in a scenario where the coverage γ is set to 0.95 and the ratio ρ equal to 1. As in the previous section, the probabilities are presented as percentages.

$m \rightarrow$	3	5	10	20	50	100	500	∞
$n = 3$	1.363	1.282	1.456	1.776	2.331	2.803	3.783	5
$n = 5$	1.282	1.012	1.015	1.252	1.765	2.251	3.386	5
$n = 10$	1.456	1.015	0.768	0.811	1.154	1.576	2.800	5
$n = 20$	1.776	1.252	0.811	0.657	0.774	1.056	2.186	5
$n = 50$	2.331	1.765	1.154	0.774	0.596	0.664	1.427	5
$n = 100$	2.803	2.251	1.576	1.056	0.664	0.577	0.980	5
$n = 500$	3.783	3.386	2.800	2.186	1.427	0.980	0.561	5
$n = \infty$	5	5	5	5	5	5	5	*

Table 2

For the Q_n^m, when ρ is set to 1, one has

$$\lim_{m\to\infty} Q_n^m(\gamma,1) = 1-\gamma \quad \text{for all } n \quad \text{and} \quad \lim_{n\to\infty} Q_n^m(\gamma,1) = 1-\gamma \quad \text{for all } m.$$

The two limits above allow for an easy intuitive explanation. Namely, with increasing m the confidence intervals generated by the sample $X_1, X_2, \ldots, X_m$ shrink to the singleton $\{\sigma_X^2\}$. The probability that a confidence interval at coverage γ, generated by a sample $Y_1, Y_2, \ldots, Y_n$, will be disjoint from this singleton is the complement of the probability that the interval will cover the number σ_X^2. Thus one arrives at the value $1-\gamma$ for the two limits above.

Besides these two intuitively clear limits it will turn out that there is a limit of the Q_n^m along the diagonal of the table given above. More precisely, the limit

$$\lim_{n\to\infty} Q_n^n(\gamma,1)$$

exists and is equal to the limit of the Q_n in [9], where the same coverage problem was dealt with when estimating the mean of a population. This result will be proved in §7.

5 Some preparatory asymptotics

In order to study the behaviour of the probabilities P_n^n and Q_n^n for large n one needs to know something about the asymptotic behaviour of the distribution functions F_n^n and that of the quantile functions q_n. The theorems in this section will prove to be useful in this. In the derivations a special convergence feature of cumulative distribution functions will be exploited several times. If, namely, a sequence $F_1, F_2, F_3, \ldots$ of cumulative distributions converges pointwise to a *continuous* distribution function F then the convergence is automatically uniform. See for example [7] for a proof of this phenomenon.

The asymptotics needed will be derived by starting from so-called infinite samples

$$X_1, X_2, X_3, \ldots$$

from a population with a standard Gaußian probability distribution. This is to say that the $X_1, X_2, X_3, \ldots$ form a statistically independent system and that they all have a standard Gaußian probability distribution. Note that for such X_i the expectation value and variance of X_i^2 is 1 and 2 respectively (see for example [7]). It follows from this that for all $n = 1, 2, 3, \ldots$ the sums Z_n, defined as

$$Z_n = \frac{X_1^2 + X_2^2 + \cdots + X_n^2 - n}{\sqrt{2n}},$$

have an expectation value equal to 0 and a variance equal to 1. In the following the cumulative distribution function of Z_n will be denoted by Φ_n and its quantile function by $\bar{q}_n$. Recall that the cumulative distribution function of the standard Gaußian distribution was convened to be denoted by Φ and its quantile function by q. In these notations one has:

Lemma 1. The Φ_n converge on $\mathbb{R}$ uniformly to Φ and the $\bar{q}_n$ on the interval (0, 1) pointwise to q.

Proof. By the central limit theorem (see [2], [7]) the Φ_n converge on $\mathbb{R}$ pointwise to Φ. The latter distribution function being continuous, this convergence is uniform. Exploiting the uniform convergence, together with the fact that Φ has a positive derivative that is locally bounded away from zero, one derives that the $\bar{q}_n$ converge pointwise to q. The necessary mathematical tools in this can be found for example in [11]. □

The following theorem describes an asymptotic feature of F-distributions by connecting them to a standard Gaußian distribution.

Theorem 2. For all $x \in \mathbb{R}$ one has

$$\lim_{n\to\infty} F_n^n\left(1 + \frac{x}{\sqrt{n}}\right) = \Phi\left(\frac{x}{2}\right).$$

The convergence is uniform in x.

Proof. Let

$$X_1, X_2, X_3, \ldots \quad \text{and} \quad Y_1, Y_2, Y_3, \ldots$$

be two independent infinite samples from a standard Gaußian distribution. Define the random variables $Z_n(X)$ and $Z_n(Y)$ as

$$Z_n(X) = \frac{X_1^2 + X_2^2 + \cdots + X_n^2 - n}{\sqrt{2n}}$$

and

$$Z_n(Y) = \frac{Y_1^2 + Y_2^2 + \cdots + Y_n^2 - n}{\sqrt{2n}}.$$

Now the $Z_n(X)$ and $Z_n(Y)$ are identically distributed; they both have Φ_n as their cumulative distribution function. Denote their probability density by φ_n. Then, in these notations, applying the law of total probability (see for example [7]), one has

$$\begin{aligned}
&F_n^n\left(1 + \sqrt{\tfrac{2}{n}}\, x\right) \\
&= \Pr\left(\frac{X_1^2 + X_2^2 + \cdots + X_n^2}{Y_1^2 + Y_2^2 + \cdots + Y_n^2} \le 1 + \sqrt{\tfrac{2}{n}}\, x\right) \\
&= \int_{-\infty}^{+\infty} \Pr\left(\frac{X_1^2 + X_2^2 + \cdots + X_n^2}{Y_1^2 + Y_2^2 + \cdots + Y_n^2} \le 1 + \sqrt{\tfrac{2}{n}}\, x \;\middle|\; Z_n(Y) = s\right) \varphi_n(s)\, \mathrm{d}s \\
&= \int_{-\infty}^{+\infty} \Pr\left(\frac{X_1^2 + X_2^2 + \cdots + X_n^2}{s\sqrt{2n} + n} \le 1 + \sqrt{\tfrac{2}{n}}\, x\right) \varphi_n(s)\, \mathrm{d}s \\
&= \int_{-\infty}^{+\infty} \Pr\left(\frac{X_1^2 + X_2^2 + \cdots + X_n^2 - n}{\sqrt{2n}} \le x + s + \sqrt{\tfrac{2}{n}}\, xs\right) \varphi_n(s)\, \mathrm{d}s \\
&= \int_{-\infty}^{+\infty} \Pr\left(Z_n(X) \le x + s + \sqrt{\tfrac{2}{n}}\, xs\right) \varphi_n(s)\, \mathrm{d}s \\
&= \int_{-\infty}^{+\infty} \Phi_n\left(x + s + \sqrt{\tfrac{2}{n}}\, xs\right) \varphi_n(s)\, \mathrm{d}s.
\end{aligned}$$

Now define the functions ε_n as

$$\varepsilon_n(x) = \int_{-\infty}^{+\infty} \left[\Phi_n\left(x + s + \sqrt{\tfrac{2}{n}}\, xs\right) - \Phi_n(x + s)\right] \varphi_n(s)\, \mathrm{d}s\,.$$

By exploiting the fact that $\Phi_n \to \Phi$ uniformly one derives that

$$\lim_{n\to\infty} \varepsilon_n(x) = 0 \qquad \text{for all } x \in \mathbb{R}\,.$$

In terms of the ε_n one may write

$$\begin{aligned}
F_n^n\left(1 + \sqrt{\tfrac{2}{n}}\, x\right) &= \varepsilon_n(x) + \int_{-\infty}^{+\infty} \Phi_n(x + s)\; \varphi_n(s)\, \mathrm{d}s \\
&= \varepsilon_n(x) + \int_{-\infty}^{+\infty} \Phi_n(x - s)\; \varphi_n(-s)\, \mathrm{d}s.
\end{aligned} \tag{4}$$

The last integral on the right side is the convolution product of the function Φ_n and the function $s \mapsto \varphi_n(-s)$, the latter being the probability density of the random variable $-Z_n(Y)$. Thus, regarding it as a function of x, the integral presents the cumulative distribution function of a random variable of type

$$Z_n(X) - Z_n(Y).$$

See for example one of the references [2], [4], [7], [14] for the underlying theory in this. Both the $Z_n(X)$ and the $Z_n(Y)$ converge in distribution to a standard Gaußian distribution. The sequences being independent, it follows that the sequence $Z_n(X) - Z_n(Y)$ converges to a Gaußian distribution with mean 0 and variance 2. Now, when taking the limit in (4) and replacing x by $x/\sqrt{2}$, one arrives at the conclusion that

$$\lim_{n\to\infty} F_n^n\left(1 + \frac{x}{\sqrt{n}}\right) = \Phi\left(\frac{x}{2}\right).$$

The limit function being continuous, the convergence above is uniform in x. □

The combination of the previous theorem with the theorem below will make the derivations in the next two sections straightforward. The theorem below describes the asymptotic behaviour of the quantile functions q_n, belonging to the family of χ^2-distributions, in terms of the quantile function q of the standard Gaußian distribution.

Theorem 3. There exists a sequence of functions $\varepsilon_n : (0, 1) \to \mathbb{R}$, converging pointwise to 0, such that

$$\frac{q_n(\eta)}{n} = 1 + \sqrt{\tfrac{2}{n}}\left[q(\eta) + \varepsilon_n(\eta)\right]$$

for all $n = 1, 2, 3\ldots$ and for all $0 < \eta < 1$. Similarly there exists a sequence of functions $\delta_n : (0, 1) \to \mathbb{R}$, converging pointwise to 0, such that

$$\frac{n}{q_n(\eta)} = 1 - \sqrt{\tfrac{2}{n}}\left[q(\eta) + \delta_n(\eta)\right]$$

for all $n = 1, 2, 3\ldots$ and for all $0 < \eta < 1$.

Proof. Let

$$X_1, X_2, X_3, \ldots$$

be an infinite sample from a standard Gaußian distribution and let the associated Z_n and $\bar{q}_n$ be as defined before. Define the functions ε_n as

$$\varepsilon_n(\eta) = \bar{q}_n(\eta) - q(\eta).$$

Then, by Lemma 1, the ε_n will converge pointwise to 0. Expressing the $\bar{q}_n$ in terms of the q_n, the $\varepsilon_n(\eta)$ may be written as:

$$\varepsilon_n(\eta) = \frac{q_n(\eta) - n}{\sqrt{2n}} - q(\eta).$$

This may be rewritten as

$$\frac{q_n(\eta)}{n} = 1 + \sqrt{\tfrac{2}{n}}\, q(\eta) + \sqrt{\tfrac{2}{n}}\, \varepsilon_n(\eta)$$

from which the first statement in the theorem follows.
The second statement can be derived from the first by *defining* the functions δ_n by

$$\frac{n}{q_n(\eta)} = 1 - \sqrt{\tfrac{2}{n}}\, \big[q(\eta) + \delta_n(\eta)\big].$$

Then the δ_n are algebraically related to the ε_n as

$$\delta_n(\eta) = \frac{\varepsilon_n(\eta) - \sqrt{\tfrac{2}{n}}\, q(\eta)\big[q(\eta) + \varepsilon_n(\eta)\big]}{1 + \sqrt{\tfrac{2}{n}}\, \big[q(\eta) + \varepsilon_n(\eta)\big]}\,.$$

For fixed η the right side converges to 0 if $n \to \infty$, thus completing the proof of the theorem. □

6 The asymptotic behaviour of the probabilities P_n^n

Exploiting the asymptotics in the previous section it is easy to describe the asymptotic behaviour of the probabilities P_n^n.

Theorem 4. For all $0 < \gamma < 1$ one has

$$\lim_{n\to\infty} P_n^n(\gamma) = 1 - 2\,\Phi\left(\frac{q\big[\frac{1}{2}(1-\gamma)\big]}{\sqrt{2}}\right).$$

Proof. By (2) in §3 the probability $P_{n+1}^{n+1}(\gamma)$ may be expressed as

$$P_{n+1}^{n+1}(\gamma) = F_n^n\left(\frac{n}{q_n(\eta_1)}\right) - F_n^n\left(\frac{n}{q_n(\eta_2)}\right) \tag{5}$$

where

$$\eta_1 = \frac{1-\gamma}{2} \quad \text{and} \quad \eta_2 = \frac{1+\gamma}{2}\,.$$

By Theorem 3 there exists a sequence of functions $\delta_n : (0, 1) \to \mathbb{R}$, converging pointwise to 0, such that

$$\frac{n}{q_n(\eta)} = 1 - \sqrt{\tfrac{2}{n}}\, \big[q(\eta) + \delta_n(\eta)\big].$$

Using this, one derives through Theorem 2 that

$$\begin{aligned}\lim_{n\to\infty} P_{n+1}^{n+1}(\gamma) &= \Phi\left(-\frac{q(\eta_1)}{\sqrt{2}}\right) - \Phi\left(-\frac{q(\eta_2)}{\sqrt{2}}\right)\\ &= \Phi\left(\frac{q(\eta_2)}{\sqrt{2}}\right) - \Phi\left(\frac{q(\eta_1)}{\sqrt{2}}\right).\end{aligned}$$

Observing that

$$\Phi\left(\frac{q(\eta_2)}{\sqrt{2}}\right) = 1 - \Phi\left(\frac{q(\eta_1)}{\sqrt{2}}\right)$$

the theorem follows. □

So, as was already announced at the end of §3, the $P_n^n(\gamma)$ are asymptotically the same as their counterparts P_n in [9], where similar coverage problems were studied in the process of estimating the mean of a population.

7 The asymptotic behaviour of the probabilities Q_n^n

The asymptotics in §5 can also be used to derive in a straightforward way the asymptotic behaviour of the probabilities $Q_n^n(\gamma)$, where $Q_n^n(\gamma)$ stands for $Q_n^n(\gamma, 1)$.

Theorem 5. For all $0 < \gamma < 1$ one has

$$\lim_{n\to\infty} Q_n^n(\gamma) = 2\,\Phi\left(\sqrt{2}\,q\left[\tfrac{1}{2}(1-\gamma)\right]\right).$$

Proof. By (3) in §4 the probability $Q_{n+1}^{n+1}(\gamma)$ may be expressed as

$$Q_{n+1}^{n+1}(\gamma) = 2\,F_n^n\left(\frac{q_n(\eta_1)}{q_n(\eta_2)}\right) \tag{6}$$

where

$$\eta_1 = \frac{1-\gamma}{2} \quad \text{and} \quad \eta_2 = \frac{1+\gamma}{2}.$$

From Theorem 3 it can be derived that there exists a sequence of functions $\theta_n : (0, 1) \to \mathbb{R}$, converging pointwise to 0, such that

$$\frac{q_n(\eta_1)}{q_n(\eta_2)} = 1 + \sqrt{\tfrac{2}{n}}\left[q(\eta_1) - q(\eta_2) + \theta_n(\gamma)\right].$$

By symmetry in the standard Gaußian distribution one has

$$q(\eta_1) = -q(\eta_2).$$

Hence one may write

$$\frac{q_n(\eta_1)}{q_n(\eta_2)} = 1 + \sqrt{\tfrac{2}{n}}\left[2\,q(\eta_1) + \theta_n(\gamma)\right].$$

Using this and Theorem 2 one derives that

$$\lim_{n\to\infty} F_n^n\left(\frac{q_n(\eta_1)}{q_n(\eta_2)}\right) = \Phi\left(\sqrt{2}\,q(\eta_1)\right).$$

In virtue of (6) this proves the statement in theorem. □

Similar to the situation in the previous section, the $Q_n^n(\gamma)$ are asymptotically the same as their counterparts Q_n in [9] where the corresponding coverage problem was studied in the estimation of the mean of a population.

8 Using interval overlap as a decision criterion

Given an $N(\mu_X, \sigma_X^2)$-distributed and an $N(\mu_Y, \sigma_Y^2)$-distributed population, the hypothesis

$$H_0 : \sigma_X = \sigma_Y$$

is sometimes tested in the following way: Two samples $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ are drawn from both populations and their corresponding 95 % confidence intervals for the variance are computed. Conclusive in the decision procedure is then whether the intervals intersect or not. If they intersect, then the hypothesis H_0 is maintained and if they are disjoint, then H_0 is rejected. When computing the confidence intervals at a coverage equal to γ, then, by Theorem 5, one arrives in this hypothesis test (for equal sample sizes) at an asymptotic significance level of

$$2\,\Phi\left(\sqrt{2}\,q\big[\tfrac{1}{2}(1-\gamma)\big]\right)$$

where, as before, q stands for the quantile function of a standard Gaußian distribution. For $\gamma = 0.95$ this leads to an asymptotic significance level of 0.5574597 % (compare this to the results in [3], [6], [9], [13]). In order to arrive in this decision procedure at an asymptotic significance level of α, the coverage γ of the two interval estimates must be adapted such as to have

$$2\,\Phi\left(\sqrt{2}\,q\big[\tfrac{1}{2}(1-\gamma)\big]\right) = \alpha.$$

Solving this equation towards γ leads to

$$\gamma = 1 - 2\,\Phi\left(\frac{q\,[\,\alpha/2\,]}{\sqrt{2}}\right).$$

If the coverage γ is set in this particular way then the asymptotic significance level of the decision procedure is equal to α. If the sample sizes are finite or unequal, however, then the significance levels will deviate from α. For arbitrary sample sizes m and n the significance level can be computed through (3), thereby taking $\rho = 1$. From now on, just to illustrate one thing and another, the asymptotic significance level α will be pinned down to 0.05. The two interval estimates must have a coverage of 0.8342315 to bring this about. Table 3, on the next page, shows the significance levels for a few (finite) values for m and n for this specific value of γ.

How does the decision procedure sketched above perform relative to Fisher's 2-sample variance test, when testing at the sample sizes and significance levels listed in the field of the table above? It seems natural to compare the two decision procedures then as to their power. As to this, denote, as before, the quotient σ_Y/σ_X by ρ. The hypothesis that is to be tested can then be formulated as

$$H_0 : \rho = 1.$$

The power of the method of disjoint intervals is presented by the probability $Q_n^m(\gamma, \rho)$. For equal sample sizes, that is for $m = n$, computations suggest that the difference in power, relative to Fisher's test, is in all cases less than 0.0001. For $m \neq n$, however, the difference in power can be considerable. For example, when taking the sample sizes $m = 5$

$m \to$	3	5	10	20	50	100	500	∞
$n = 3$	6.716	6.639	7.418	8.711	10.65	12.00	14.28	16.58
$n = 5$	6.639	5.972	6.151	7.115	8.989	10.50	13.37	16.58
$n = 10$	7.418	6.151	5.457	5.724	7.104	8.580	12.00	16.58
$n = 20$	8.711	7.115	5.724	5.221	5.773	6.903	10.46	16.58
$n = 50$	10.65	8.989	7.104	5.773	5.086	5.413	8.273	16.58
$n = 100$	12.00	10.50	8.580	6.903	5.413	5.043	6.749	16.58
$n = 500$	14.28	13.37	12.00	10.46	8.273	6.749	5.001	16.58
$n = \infty$	16.58	16.58	16.58	16.58	16.58	16.58	16.58	$*$

Table 3

and $n = 10$, the power of the method of disjoint intervals in $\rho = 3$ exceeds the power in Fisher's test by more than 0.04. This particular evaluation shows that, when fixing some significance level, Fisher's 2-sample variance test does not automatically realize maximum power. Fisher's test is in some cases outperformed by the method of disjoint intervals. In other cases, however, it is the other way round. Fisher's test is an example of a maximum likelihood test. It is known that such tests do not automatically maximize power at fixed significance levels. See for example [7] for more details in this. As a closing remark, in the above Fisher's test was carried out in the way it is carried out in the powerful open-source statistical package R (see [10]). That is to say, the left and right critical regions in the test are taken to be of equal probabilistic size. Otherwise formulated, in Fisher's variance test the two-sided p-values are chosen to be twice the right-sided p-values.

References

[1] Barr, D.R.: Using confidence intervals to test hypotheses, *Journal of Quality Technology* 1 (1969), pp. 256–258.

[2] Bauer, H.: *Probability Theory*, translated from the originally German version, Walter de Gruyter Verlag, Berlin (1996).

[3] Goldstein, H.; Healy, M.J.R.: The Graphical Presentation of a Collection of Means, *Journal of the Royal Statistical Association*, series A, 158, Part 1, (1995), pp. 175–177.

[4] Grimmett, G.R.; Stirzaker, D.R.: *Probability theory and random processes*, 3rd ed., Oxford University Press (2001).

[5] Nelson, L.S.: Evaluating overlapping confidence intervals, *Journal of Quality Technology* 21 (1989), pp. 140–141.

[6] Payton, M.E.; Greenstone, M.H.; Schenker, N.: Overlapping confidence intervals or standard error intervals: what do they mean in terms of statistical significance?, *Journal of Insect Science* 3, (2003), pp. 34–39. Available at www.insectscience.org/3.34.

[7] Pestman, W.R.: *Mathematical Statistics*, 2nd edition, Walter de Gruyter Verlag, Berlin (2009).

[8] Pestman, W.R.: Teaching the independence of $\bar{X}$ and S^2, *Elemente der Mathematik* 53 (1998), pp. 107–111.

[9] Pestman, W.R.: Two elementary statistical coverage problems, *Elemente der Mathematik* 66 (2011), pp. 30–37.

[10] R Development Core Team: *R: A Language and Environment for Statistical Computing*, http: www.R-project.org.

[11] Rudin, W.: *Principles of Mathematical Analysis*, McGraw-Hill, New York (1976).

[12] Ryan, G.W.; Leadbetter, S.D.: On the misuse of confidence intervals for two means in testing for the significance of the difference between the means, *Journal of Modern Applied Statistical Methods* 1 (2002), pp. 473–478.

[13] Schenker, N.; Gentleman, J.F.: On judging the significance of differences by examining the overlap between confidence intervals, *Am. Statistician* 55 (2001), pp. 182–186.

[14] Schwartz, L.: *Mathematics for the Physical Sciences*, Dover Publications, New York (2008).

[15] Ventsel, H.: *Théorie des Probabilités*, Éditions Mir, Moscou (1973).

[16] Zar, J.H.: *Biostatistical Analysis*, Pearson Prentice Hall (2010).

Wiebe R. Pestman
Universidade Federal de Santa Catarina
Departamento de Matematica
Florianopolis SC
Brasil
e-mail: w.r.pestman@tele2.nl