Max–Planck–Institut für biologische Kybernetik
Max Planck Institute for Biological Cybernetics

MAX-PLANCK-GESELLSCHAFT

Technical Report No. TR-133

Confidence Sets for Ratios: A Purely Geometric Approach To Fieller's Theorem

Ulrike von Luxburg,[1] Volker H. Franz[2]

December 2004

[1] Department for Empirical Inference, email: ulrike.luxburg@tuebingen.mpg.de
[2] Justus–Liebig–Universität Giessen, FB 06 / Abt. Allgemeine Psychologie, Otto–Behaghel–Strasse 10F, 35394 Giessen, Germany, email: volker.franz@psychol.uni-giessen.de

This report is available in PDF–format via anonymous ftp at ftp://ftp.kyb.tuebingen.mpg.de/pub/mpi-memos/pdf/TRfieller04.pdf. The complete series of Technical Reports is documented at: http://www.kyb.tuebingen.mpg.de/techreports.html

Confidence Sets for Ratios: A Purely Geometric Approach To Fieller's Theorem

Ulrike von Luxburg, Volker H. Franz

Abstract. We present a simple, geometric method to construct Fieller's exact confidence sets for ratios of jointly normally distributed random variables. Contrary to previous geometric approaches in the literature, our method is valid in the general case where both sample mean and covariance are unknown. Moreover, not only the construction but also its proof are purely geometric and elementary, thus giving intuition into the nature of the confidence sets.

1 Introduction

In many practical applications we encounter the problem of estimating the ratio of two random variables X and Y. This could, for example, be the case if we want to know how large one quantity is relative to the other, or if we want to estimate at which position a regression line intersects the abscissa (cf. Marsaglia 1965). While it is straightforward to construct an estimator for $E(Y)/E(X)$ by dividing the two sample means of X and Y, it is not obvious how confidence regions for this estimator can be constructed. In the case where X and Y are jointly normally distributed, an exact solution to this problem has been derived by Fieller (1932, 1954). But in practice, his results are seldomly used, and they are not very well-known among non-statisticians. Perhaps the main reason why Fieller's results are so unpopular among practitioners is that his confidence regions do not look like "normal" confidence intervals and are often perceived as counter-intuitive. In benign cases they form an interval which is not symmetric around the estimator, while in worse cases the confidence region consists of two disjoint unbounded intervals, or even of the whole real line. This raised the suspicion that confidence regions for ratios are a highly complicated issue, and that Fieller's solutions lack intuition and are difficult to interpret. As a consequence, many ad–hoc heuristics have been used. For a discussion and empirical comparison see Franz (submitted 2004).

There have been several approaches to simplify Fieller's proofs. Especially remarkable are the ones which rely on geometric arguments, as they might lead to more intuition about the constructed confidence sets. Milliken (1982) attempted a geometric proof for Fieller's result in the case where X and Y are independent normally distributed random variables. Unfortunately, his proof contained an error which led him to the wrong conclusion that Fieller's confidence regions were too conservative. Later, his proof was corrected and simplified by Guiard (1989). He considers the case that X and Y are jointly normally distributed according to $(X,Y) \sim N(\mu, \sigma^2 V)$, where the mean μ and the scale σ^2 of the covariance are unknown, but the covariance matrix V is known. Guiard presents a geometric construction of confidence regions, and then shows by an elegant comparison to a likelihood ratio test that the constructed regions are exact and coincide with Fieller's solution. The drawback of his proof is that it only works in the case where the covariance matrix V is known, which is usually not the case in practice. Moreover, although the confidence sets are constructed by a geometric procedure, Guiard's proof relies on properties of the likelihood ratio test and does not give geometric insights why the construction is correct.

In this article we derive a simple geometric construction for confidence intervals for ratios. Contrarily to Milliken (1982) and Guiard (1989), this construction is valid in the general setting where $(X,Y) \sim N(\mu, C)$ with unknown mean μ and unknown covariance matrix C. Our construction is related to the construction in Guiard (1989), but its derivation is new and completely different. Our proofs are rather elementary, purely geometric, and lead to insights why our construction is correct. The goal in this paper is to make these geometric insights available to a broad audience and show that Fieller's confidence regions are indeed a very natural solution.

1.1 Definitions and Notation

We will always consider the following situation. We are given a sample of n pairs $Z_i := (X_i, Y_i)_{i=1,\dots,n}$ drawn independently according to the joint 2-dimensional normal distribution $N(\mu, C)$ with mean $\mu = (\mu_1, \mu_2)$ and covariance matrix C. We assume that both μ and C are unknown. Our goal will be to estimate the ratio $\rho := \mu_2/\mu_1$ and construct confidence intervals for this estimate.

To estimate the unknown mean and the covariance matrix of the samples we will use the standard estimators: the means are estimated by

$$\hat{\mu}_1 := \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } \hat{\mu}_2 := \frac{1}{n}\sum_{i=1}^{n} Y_i, \tag{1}$$

and the estimated covariance matrix $\hat{C}$ has the entries

$$\hat{c}_{11} := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu}_1)^2 \text{ and } \hat{c}_{22} := \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \hat{\mu}_2)^2 \tag{2}$$

$$\hat{c}_{12} := \hat{c}_{21} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\mu}_1)(Y_i - \hat{\mu}_2). \tag{3}$$

As estimator for the ratio $\rho = \mu_2/\mu_1$ we use

$$\hat{\rho} := \frac{\hat{\mu}_2}{\hat{\mu}_1}. \tag{4}$$

Our goal will be to construct confidence sets for ρ. In general, for $\alpha \in]0,1[$, a confidence set (or confidence region) of level $1-\alpha$ for a quantity ρ is defined to be a set I constructed from the sample such that $P(\rho \in I) \geq 1-\alpha$. Remember that the random quantity in this statement is the data-dependent region I and not the parameter ρ. If $P(\rho \in I) \geq 1-\alpha$ always holds with equality, then the confidence set I is called exact, otherwise it is called conservative. For more discussion about confidence sets we refer to Chapter 20 of Kendall and Stuart (1961) or Section 5.2 of Schervish (1995).

Before we continue we would like to point out several things. Note that we want to estimate $E(Y)/E(X)$ and not $E(Y/X)$. In fact, the latter quantity does not exist. In case where X and Y are uncorrelated this is easy to see. As X is normally distributed, the density of X is positive in every ε-neighborhood of 0. Together with the fact that the integral $\int_{-\varepsilon}^{\varepsilon} 1/x \; dx$ is infinite this shows that $E(1/X)$ and $E(Y/X)$ do not exist. Similar statements are true in the correlated case. The observation that $E(Y/X)$ does not exist for normally distributed random variables, together with the fact that the estimators $\hat{\mu}_1$ and $\hat{\mu}_2$ are normally distributed, also shows that the estimator $\hat{\rho}$ (as well as any other estimator for ρ) cannot be unbiased. Its expectation $E(\hat{\rho}) = E(\hat{\mu}_2/\hat{\mu}_1)$ simply does not exist. For more discussion about the distribution of the ratio of two normal distributions we refer to Marsaglia (1965), Hinkley (1969, 1970) and references therein.

2 Geometric construction of exact confidence regions

In this section we want to explain how confidence sets for the ratio ρ can be constructed geometrically. Let us start with some simple geometric observations, most of which were already present in Fieller (1954). For given $\mu = (\mu_1, \mu_2) \in \mathbb{R}^2$, the ratio $\rho = \mu_2/\mu_1$ can be depicted as the slope of the line in the two-dimensional plane which passes both through the origin and the point (μ_1, μ_2). Similarly, the estimated ratio $\hat{\rho}$ is given as the slope of the line through the origin and the point $\hat{\mu} = (\hat{\mu}_1, \hat{\mu}_2)$ (cf. Figure 1). Assume that we are given a confidence interval $I = [l, u] \subset \mathbb{R}$ that contains the estimator $\hat{\rho}$. The lower and upper limits of this interval correspond to the slopes of the two lines passing through the origin and the points $(1, l)$ and $(1, u)$, respectively. The lines inside the wedge W enclosed by the two lines exactly correspond to the ratios inside the interval I. This also works the other way round: given a wedge containing the line with slope $\hat{\rho}$, the interval $[l, u]$ can be reconstructed from the wedge as the intersection of the wedge with the line $x = 1$ (cf. Figure 1).

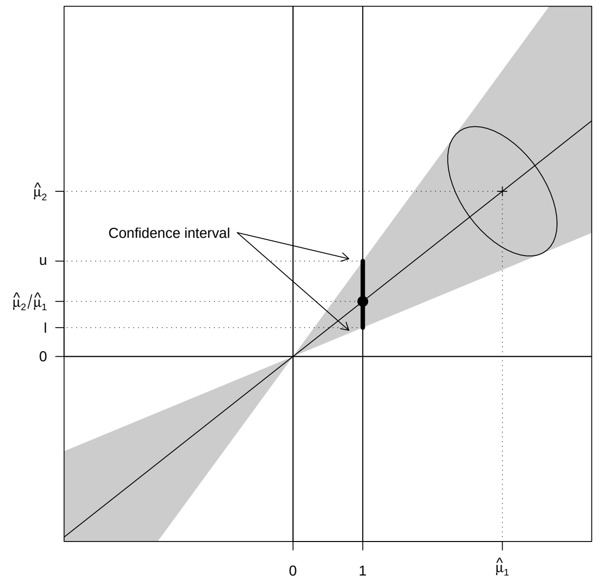

Figure 1: Geometric principles. The ratio $\hat{\mu}_2/\hat{\mu}_1$ can be depicted as the slope of the line through the points $(0, 0)$ and $(\hat{\mu}_1, \hat{\mu}_2)$. The ratios inside an interval $[l, u]$ correspond to the slopes of all lines in the wedge spanned by the lines with slopes l and u. For a given wedge, the corresponding interval $[l, u]$ can be obtained by intersecting the wedge with the line $x = 1$. Later, an appropriate wedge will be constructed by fitting it around a certain ellipse centered at $(\hat{\mu}_1, \hat{\mu}_2)$.

In the following we want to construct an appropriate wedge containing $\hat{\mu}$ such that the region obtained by intersection with the line $x = 1$ yields an exact confidence region for ρ of level $1 - \alpha$. This wedge will be constructed as the smallest wedge containing a certain ellipse around the estimated mean $(\hat{\mu}_1, \hat{\mu}_2)$. We will see that we have to distinguish between three different cases, depending on the position of the ellipse. The first case will occur if the denominator $\hat{\mu}_1$ is significantly different from 0. Here the ellipse lies completely on one side of the y-axis. The corresponding wedge does not contain the y-axis, and the confidence set obtained by intersecting the wedge with the line $x = 1$ will be a bounded interval (see the left two panels of Figure 2). We will call this case the *"bounded case"*.

A not so benign situation will occur if the the denominator is not significantly different from 0. We will see that in this case, the ellipse intersects the y-axis, and it now depends on the actual position of the ellipse how to proceed. Firstly, assume that the ellipse does not contain the origin. In this case we can proceed as above and construct the wedge given by the two tangents of the ellipse through the origin. The confidence region which is given as the intersection between the wedge and the line $x = 1$ will then be unbounded, but have a small hole in the middle (see third panel of Figure 2). We will call the corresponding region *"exclusive unbounded"*. Later we will see that this case occurs if the denominator is "close to 0" (i.e., not significantly different from 0), but the numerator is far from 0. Intuitively, the form of the confidence set can be explained by considering what happens if we divide a fairly large number by a number close to 0. The magnitude of the ratio can get arbitrary large, and as we cannot be sure about the sign of the denominator, we also do not know the sign of the ratio. Thus it makes sense that regions of the form $]-\infty, c_1]$ and $[c_2, \infty[$ are part of the confidence region. The fact that there is an excluded "hole" in the confidence set is related to the only thing we are relatively certain about: the only way to end up in the excluded part of the real line would be either to have a very small numerator or a very large denominator, which we both believe to be rather unlikely because of the assumptions in this case.

The last case occurs if the origin is contained in the ellipse. In this case it is impossible to construct tangents to the ellipse which go through the origin, which means that our construction for the wedge will break down (see right panel of Figure 2). In this case we will choose the unbounded confidence set $]-\infty, \infty[$, and we will call this the *"completely unbounded case"*. Intuitively this makes sense, as this case occurs if both numerator and denominator are not significantly different from 0. Thus we believe that the ratio ρ is close to $0/0$, which can take

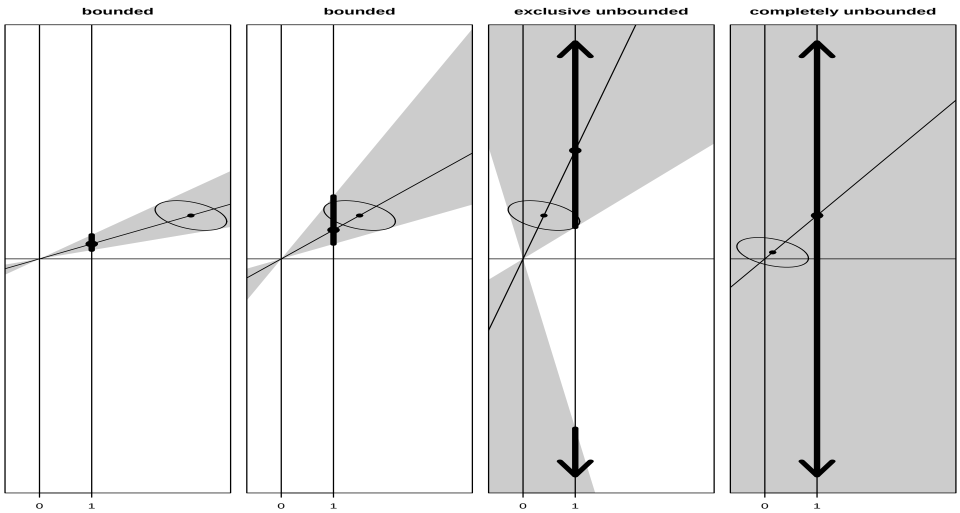

Figure 2: The three cases in the construction of the confidence set R (for figure labels please consider Figure 1): the bounded case where the ellipse does not intersect the y-axis (left two panels); the exclusive unbounded case, where the ellipse intersects the y-axis but does not contain the origin (third panel); the completely unbounded case, where the ellipse contains the origin (right panel). Note also that even in the bounded case, the estimator $\hat{\mu}_2/\hat{\mu}_1$ is usually not the geometric center of the confidence set (second panel).

arbitrary real values.

To explain the construction of the ellipse in detail we need to introduce some notation concerning covariance ellipses. Let $C \in I\!R^{2\times 2}$ a covariance matrix (i.e., positive definite and symmetric) with eigenvectors $v_1, v_2 \in I\!R^2$ and eigenvalues $\lambda_1, \lambda_2 \in I\!R$. Consider the ellipse centered at some point $\mu \in I\!R^2$ such that its principal axes have the directions of v_1, v_2 and have lengths $q\sqrt{\lambda_1}$ and $q\sqrt{\lambda_2}$ for some $q > 0$. We denote this ellipse by $E(C, \mu, q)$ and call it the covariance ellipse according to C centered at μ and scaled with parameter q. This ellipse can also be described as the set of points $z \in I\!R^2$ which satisfy the ellipse equation $(z-\mu)'C^{-1}(z-\mu) = q^2$.

Now we can explain how we can construct exact confidence regions for ρ. The proof that this construction is correct will be postponed to Section 4.

Geometric construction of exact confidence regions R of level $(1-\alpha)$ for ρ

1. Estimate the means $\hat{\mu}_1$ and $\hat{\mu}_2$ according to Equation (1), the covariance matrix $\hat{C}$ according to Equation (2).
2. Define the real number q to be half the length of a $(1-\alpha)$-quantile of the Student-t distribution with $n-1$ degrees of freedom.
3. In the two-dimensional plane, plot the ellipse $E = E(\hat{C}, \hat{\mu}, q)$ centered at the estimated joint mean $\hat{\mu} = (\hat{\mu}_1, \hat{\mu}_2)$, with shape according to the estimated covariance matrix $\hat{C}$, and scaled by the number q computed in the step before.
4. Depending on the position of the ellipse, distinguish between the following cases (see Figure 2).
   (a) If $(0,0) \notin E$, construct the two tangents to E which go through the origin $(0,0)$ and let W be the wedge enclosed by those tangents. Define the region R as the intersection of W with the line $x = 1$. Depending on whether the y-axis is contained in W or not, this results in an exclusive unbounded or a bounded confidence region.
   (b) If $(0,0) \in E$, choose the confidence region as $R =]-\infty, \infty[$ (completely unbounded case).

There are several things to note about this construction. The first thing is that usually the estimator $\hat{\rho}$ is not in the geometric center of the region R. Even in the bounded case, the closer the ellipse gets to the y-axis, the more $\hat{\rho}$ deviates from the center of the confidence interval (see first two panels of Figure 2). But it can be seen that the region R is symmetric in the following sense:

$$P(\rho \in R \text{ and } \rho \leq \hat{\rho}) = P(\rho \in R \text{ and } \rho \geq \hat{\rho}) = \frac{1}{2}(1-\alpha). \quad (5)$$

This means that the region R is central: the probability that we estimated too high and the true value is in the lower part of the region equals the probability that we estimated too low and the true value is in the upper part of the region (cf. Section 20.7. of Kendall and Stuart, 1961).

Secondly we note that by decreasing or increasing the confidence level α we can switch between the three cases "bounded", "exclusive unbounded", and "completely unbounded". If we choose α small enough the ellipse E will not intersect the y-axis, and consequently we are in the benign case of bounded confidence regions. On the other hand we can always choose α so large that the ellipse will cover the origin and we get a completely unbounded confidence region.

This observation leads to one aspect of the construction of R which might be confusing at first sight. As an example, assume that we are given a sample such that the ellipse E of confidence level 95% just touches the origin. In this case, the corresponding confidence set R_{95} is unbounded. On the other hand, the ellipse for confidence level 94.99% is a bit smaller and does not touch the origin, and the corresponding confidence region $R_{94.99}$ is exclusive unbounded. Now one is tempted to think that that $P(\rho \in R_{94.99}) = 94.99\%$ but $P(\rho \in R_{95}) = P(\rho \in \mathbb{R}) = 1$. It seems that "we have lost" 5% of probability here. The solution of this apparent paradox lies (as so often) in the definition of a confidence region: it is not the parameter ρ which is random, but the set R. For a *particular sample* and its confidence region R it is *not true* that $P(\rho \in R) = 95\%$. For fixed sample, this statement does not even make sense as ρ is not a random variable. The statement $P(\rho \in R) = 95\%$ only holds *in average* over all samples. It is by definition impossible to make any statement about a particular sample.

3 Generalizations

The geometric principle that a confidence set can be constructed as the intersection of a wedge with the line $x = 1$ can also be used to derive other confidence sets for ρ. For example, we can replace the ellipse E from the construction above by a convex two-dimensional $(1-\alpha)$-confidence set $M \subset \mathbb{R}^2$ for the two-variate joint mean $\mu \in \mathbb{R}^2$, that is a set such that $P(\mu \in M) = 1 - \alpha$. Then, as above we can construct the wedge W around M which is given by the two enclosing tangents and choose a confidence region S by intersecting the wedge with the line $x = 1$, distinguishing between the same three cases as above. By construction we then have that

$$P(\rho \in S) = P(\mu \in W) \geq P(\mu \in M) = 1 - \alpha$$

Because of the inequality in this statement, the so constructed confidence set S will be conservative in general. Formally, the construction is the following:

General geometric construction of conservative confidence regions S of level $(1-\alpha)$ for $\boldsymbol{\rho}$

1. Estimate the means $\hat{\mu}_1$ and $\hat{\mu}_2$ according to Equation (1), the covariance matrix $\hat{C}$ according to Equation (2).
2. In the two-dimensional plane, construct a convex confidence region M of level $(1-\alpha)$ for the joint mean (μ_1, μ_2).
3. (a) If $(0,0) \notin M$, construct the two tangents to M which go through the origin $(0,0)$ and let W be the wedge enclosed by those tangents (such that $M \subset W$). Define the confidence region S as the intersection of W with the line $x = 1$. Depending on whether the y-axis is contained in W or not this results in an exclusive unbounded or a bounded confidence region
   (b) $(0,0) \in M$, choose the confidence region as $S =]-\infty, \infty[$.

There are several obvious ways to choose the set M. For example, a two-dimensional confidence set for the joint mean μ is given by the ellipse $E_2(\hat{\mu}, \hat{C}, r)$ centered at $\hat{\mu}$ with shape according to the estimated covariance $\hat{C}$, where the scaling parameter r is given as half the length of the $(1-\alpha)$-quantile of the Hotelling T^2 distribution with dimension 2 and $n-1$ degrees of freedom (cf. Section 3.3.3 of Rencher, 1998). Using the ellipse E_2 in the general construction of the confidence sets S leads to conservative confidence sets S of level $(1-\alpha)$ (see Figure 3).

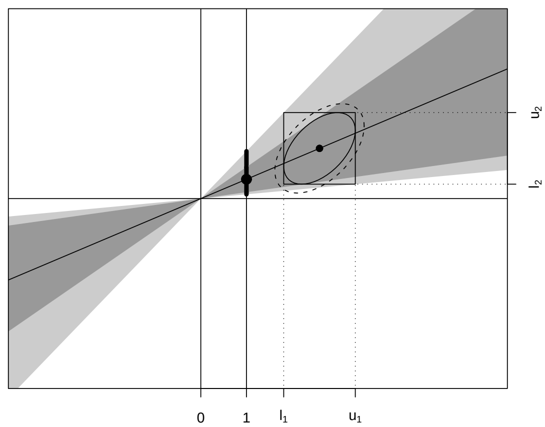

Figure 3: Geometric construction of conservative confidence regions. The small ellipse is the one used in the exact construction of Section 2 (called E in the text). The larger, dashed ellipse (called E_2 in the text) is a two-dimensional confidence region for the joint mean μ of level $1-\alpha$. The rectangle (called A in the text) is a two-dimensional conservative confidence set for μ of level $1-2\alpha$. As before, the confidence sets for ρ are obtained by intersecting the wedge with the line $x = 1$. Here we show the confidence set for ρ which is constructed from the rectangle.

We can also choose M to be a rectangle. Let the intervals $I_1 := [l_1, u_1]$ and $I_2 := [l_2, u_2]$ be the standard confidence intervals for the one-dimensional means μ_1 and μ_2 of level $1-\alpha$, as given by the quantile of the t-distribution with $n-1$ degrees of freedom. By definition we then have that $P(\mu_1 \in I_1) = P(\mu_2 \in I_2) = 1-\alpha$. Now let A be the axis-parallel rectangle $I_1 \times I_2$, centered at the estimated mean $(\hat{\mu}_1, \hat{\mu}_2)$ (see Figure 3). Let S the confidence set obtained from this rectangle according to the construction above. By construction we now have

$$
\begin{aligned}
P(\rho \in S) &= P(\mu \in W) \\
&\geq P(\mu \in A) \\
&= P(\mu_1 \in I_1 \text{ and } \mu_2 \in I_2) \\
&= 1 - P(\mu_1 \notin I_1 \text{ or } \mu_2 \notin I_2) \\
&\geq 1 - (P(\mu_1 \notin I_1) + P(\mu_2 \notin I_2)) \\
&= 1 - 2\alpha.
\end{aligned}
$$

This shows that the rectangle construction leads to a conservative confidence set of level $(1-2\alpha)$. It can also be seen easily that the set S coincides with the set obtained by "dividing" the one-dimensional confidence intervals I_2 by I_1, namely

$$S = I_2/I_1 := \{\frac{y}{x};\ y \in I_2, x \in I_1\}.$$

Note that the construction for conservative confidence sets presented in this section is in fact a fairly general principle which can also be applied to situations where the distributions of the random variables are not normal. For example, in case of unknown distributions this principle can be used to construct bootstrap confidence sets for ratios of random variables.

4 Proofs

In this section we want to prove that the construction for the exact confidence regions R as introduced in Section 2 is correct.

Theorem 1 (R is an exact confidence set for ρ) *Let $(X, Y) \sim N(\mu, C)$ with unknown μ and C, and let R the regions constructed according to the procedure in Section 2. Then R is an exact confidence region of level $1 - \alpha$ for ρ, that is for all μ and C we have $P(\rho \in R) = 1 - \alpha$.*

The idea of the proof is very simple. Consider an orthogonal projection π_a of the two-dimensional space on the subspace spanned by some unit vector $a \in I\!R^2$. It maps the ellipse to an interval, the estimated joint mean $\hat{\mu}$ to the center of this interval, and the true joint mean μ to some other point. According to our construction we chose the width of the ellipse exactly such that $\pi_a(E)$ is a confidence interval of level $1 - \alpha$ for $\pi_a(\mu)$. This holds for the projections for all a, so in particular for the projection $\pi_{\rho_\perp}$ on the line which is perpendicular to the line with slope ρ. For this projection it is easy to see (Figure 4) that $\pi_{\rho_\perp}(\mu) \in \pi_{\rho_\perp}(E)$ iff $\rho \in R$, and by construction this happens with probability $1 - \alpha$. Now let us give the details.

Proof. In the whole proof we denote by $a = (a_1, a_2) \in I\!R^2$ an arbitrary unit vector and by

$$\pi_a : I\!R^2 \to I\!R, \; x \mapsto a'x = a_1 x_1 + a_2 x_2$$

the projection of the two-dimensional space on the one-dimensional subspace spanned by a. We denote by $U := \pi_a(X, Y)$ the projection of the joint random variable (X, Y) on a. It is well-known that if (X, Y) is distributed according to $N(\mu, C)$, then U is distributed according to $N(a'\mu, a'Ca)$. The independent sample points $(X_i, Y_i)_{i=1,\ldots n}$ are mapped by π_a to independent sample points $(U_i)_{i=1,\ldots,n}$. We now subdivide the proof in four easy steps. For illustration we refer to Figure 4.

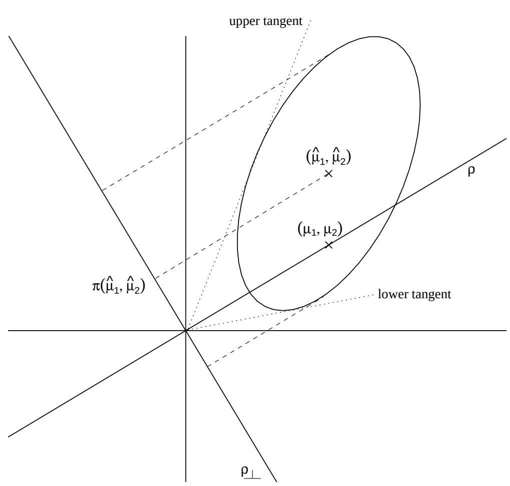

Figure 4: Projection of the ellipse E on the subspace spanned by $\rho_\perp$ (see proof of Theorem 1).

1. The length of the interval $I := \pi_a(E)$ is $2(a'\hat{C}a)^{1/2}$. This follows from the two well-known facts that the ellipse E depicts the standard deviations of a $N(\hat{\mu}, \hat{C})$-distributed random variable in the different directions a, and that the latter are given by the formula $(a'\hat{C}a)^{1/2}$.

2. It makes no difference whether we first project the sample $(X_i, Y_i)_{i=1,\ldots,n}$ on a and then estimate mean and variance of $(U_i)_{i=1,\ldots,n}$ or whether we first estimate joint mean and covariance of the sample $(X_i, Y_i)_{i=1,\ldots,n}$ and then project the results on a.

Denote by $\nu := E(U)$ the true mean of U, by $\hat{\nu} := \frac{1}{n}\sum_{i=1}^{n} U_i$ the estimated sample mean, and by $\hat{\sigma}_a^2 := \frac{1}{n-1}\sum_{i=1}^{n}(U_i - \hat{\nu})^2$ the estimated sample variance of $U_1, ..., U_n$. By linearity of the projection and the expectation it is clear that $\nu = \pi_a(\mu)$ and $\hat{\nu} = \pi_a(\hat{\mu})$. Moreover we have that $\hat{\sigma}_a^2 = a'\hat{C}a$, which can be seen as follows:

$$\begin{aligned}
\hat{\sigma}_a^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(U_i - \hat{\nu})^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}(\pi_a(X_i, Y_i) - \pi_a(\hat{\mu}))^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}(a_1X_i + a_2Y_i - a_1\hat{\mu}_1 - a_2\hat{\mu}_2)^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n} a_1^2(X_i - \hat{\mu}_1)^2 + a_2^2(Y_i - \hat{\mu}_2)^2 + 2a_1a_2(X_i - \hat{\mu}_1)(Y_i - \hat{\mu}_2) \\
&= a_1^2\hat{c}_{11} + a_2^2\hat{c}_{22} + a_1a_2\hat{c}_{12} + a_2a_1\hat{c}_{21} \\
&= a'\hat{C}a
\end{aligned}$$

3. $\pi_a(E)$ is a $(1-\alpha)$-confidence interval for ν. In the construction of the region R, the scaling factor q has been chosen such that it coincides with half the length of a $t_{n-1}(1-\alpha)$-confidence interval. As U is normally distributed and by steps (2) and (1) of the proof we have

$$\begin{aligned}
1-\alpha &= P(\nu \in [\hat{\nu} - q\hat{\sigma}_a, \hat{\nu} + q\hat{\sigma}_a]) \\
&= P(\pi_a(\mu) \in [\pi_a(\hat{\mu}) - q(a'\hat{C}a)^{1/2},\ \pi_a(\hat{\mu}) + q(a'\hat{C}a)^{1/2}]) \\
&= P(\pi_a(\mu) \in \pi_a(E)).
\end{aligned}$$

4. Projection on $\rho_\perp$. So far we have seen that for all projections π_a we have that

$$P(\pi_a(\mu) \in \pi_a(E)) = 1 - \alpha. \tag{6}$$

Now consider the special case where we project on the vector $\rho_\perp = (\rho/\sqrt{1+\rho^2}, -1/\sqrt{1+\rho^2})$ which is perpendicular to the line through the origin and the true mean μ. Our goal is now to prove that

$$\pi_{\rho_\perp}(\mu) \in \pi_{\rho_\perp}(E) \iff \rho \in R,$$

as this together with Equation (6) proves that R is an exact confidence region. As in the construction of R we consider two cases. In the first case the origin is not contained in the ellipse E. In this case we can construct the wedge W as described in the construction of R. Denoting the line with slope ρ by L_ρ, we now have the following geometric equivalences (see Figure 4):

$$\pi_{\rho_\perp}(\mu) \in \pi_{\rho_\perp}(E) \iff 0 \in \pi_{\rho_\perp}(E) \iff E \cap L_\rho \neq \emptyset \iff L_\rho \subset W \iff \rho \in R.$$

In the second case, the origin is contained in the ellipse E. Under this assumption it is clear that $\pi_{\rho_\perp}(\mu) = 0$ is always contained in $\pi_{\rho_\perp}(E)$. On the other hand, in this case the region R coincides with $]-\infty, \infty[$ by construction. Thus we also have that always $\rho \in R$ is true. Consequently, in this case we have that $\pi_{\rho_\perp}(\mu) \in \pi_{\rho_\perp}(E) \iff \rho \in R$. This completes the proof. ☺

Theorem 1 shows that the confidence regions R obtained by our construction are exact confidence regions. Now we want to compare our solution to the classic confidence sets constructed by (Fieller, 1954). To this end let us first state Fieller's result according to Subsection 4, p. 176-177 of (Fieller, 1954) (note that there is a typo in Formula (9) of Fieller (1954); in the first numerator the quantity x^2v_{xy} should be replaced by x^2v_{yy}). We reformulate his result in our notation:

Theorem 2 (Fieller's confidence regions) *Compute the quantities*

$$q^2_{\textit{exclusive}} := \frac{\hat{\mu}_1^2}{\hat{c}_{11}} \quad \textit{and} \quad q^2_{\textit{complete}} := \frac{\hat{\mu}_2^2\hat{c}_{11} - 2\hat{\mu}_1\hat{\mu}_2\hat{c}_{12} + \hat{\mu}_1^2\hat{c}_{22}}{\hat{c}_{11}\hat{c}_{22} - \hat{c}_{12}^2}$$

and the two numbers

$$l_{1,2} = \frac{(\hat{\mu}_1\hat{\mu}_2 - q^2\hat{c}_{12}) \mp \sqrt{(\hat{\mu}_1\hat{\mu}_2 - q^2\hat{c}_{12})^2 - (\hat{\mu}_1^2 - q^2\hat{c}_{11})(\hat{\mu}_2^2 - q^2\hat{c}_{22})}}{\hat{\mu}_1^2 - q^2\hat{c}_{11}}$$

with q as in the definition of the confidence regions R. Define the quantity q as in the geometric construction of R. Then exact confidence sets F of level $1-\alpha$ for the ratio ρ can be defined as follows:

$$F = \begin{cases}]-\infty, \infty[& \text{if } q^2_{\text{complete}} \leq q^2 \quad \text{(completely unbounded case)} \\ ]-\infty, l_1] \quad \cup \quad [l_2, \infty[& \text{if } q^2_{\text{exclusive}} < q^2 < q^2_{\text{complete}} \quad \text{(exclusive unbounded case)} \\ [l_1, l_2] & \text{otherwise} \quad \text{(bounded case)} \end{cases}$$

Now we are given two ways to compute exact confidence sets for ρ: our confidence sets R and Fieller's confidence sets F. A priori it is not clear that both sets coincide, as confidence sets are not necessarily unique. But now we will see that the two constructions indeed coincide.

Theorem 3 (R coincides with Fieller's F) *The confidence regions R constructed in Section 2 coincide with Fieller's confidence regions F.*

Proof. First we want to show that the three cases of Fieller's theorem coincide with our three cases. According to our construction, the completely unbounded case occurs if the origin is contained in the ellipse E. By the ellipse equation this is the case if

$$(0 - \hat{\mu})'\hat{C}^{-1}(0 - \hat{\mu}) \leq q^2. \tag{7}$$

Using the fact that the inverse of the symmetric positive definite matrix $\hat{C}$ can be computed by the formula

$$C^{-1} = \frac{1}{\hat{c}_{11}\hat{c}_{22} - \hat{c}_{12}} \begin{pmatrix} \hat{c}_{22} & -\hat{c}_{12} \\ -\hat{c}_{12} & \hat{c}_{11} \end{pmatrix}$$

it is easy to check that Equation (7) is equivalent to the condition $q^2_{\text{complete}} \leq q^2$ of Fieller's theorem.
In our construction of R, the exclusive unbounded case occurs if the ellipse intersects the y-axis. This is equivalent to the condition that 0 is contained in the projection of the ellipse on the x-axis, that is $0 \in [\hat{\mu}_1 - q\sqrt{\hat{c}_{11}}, \hat{\mu}_1 + q\sqrt{\hat{c}_{11}}]$. This condition is equivalent to the condition $(\hat{\mu}_1 - q\sqrt{\hat{c}_{11}})(\hat{\mu}_1 + q\sqrt{\hat{c}_{11}}) \leq 0$, which coincides with the condition $q^2_{\text{exclusive}} < q^2$ of Fieller. Thus our three cases coincide with the ones of Fieller.

Now it remains to prove that in the exclusive bounded and the bounded case, the limits l_1, l_2 of the Fieller's theorem coincide with the slopes of the two tangents to the ellipse. To compute the slopes of the two tangents one can solve the optimization problems

$$\min_{x,y \in I\!R} \frac{y}{x} \quad \text{subject to} \quad \begin{pmatrix} x - \hat{\mu}_1 \\ y - \hat{\mu}_2 \end{pmatrix}' \hat{C}^{-1} \begin{pmatrix} x - \hat{\mu}_1 \\ y - \hat{\mu}_2 \end{pmatrix} = q^2$$

and

$$\max_{x,y \in I\!R} \frac{y}{x} \quad \text{subject to} \quad \begin{pmatrix} x - \hat{\mu}_1 \\ y - \hat{\mu}_2 \end{pmatrix}' \hat{C}^{-1} \begin{pmatrix} x - \hat{\mu}_1 \\ y - \hat{\mu}_2 \end{pmatrix} = q^2.$$

By some tedious calculations it can be verified that the slopes of the two tangents are given by the two expressions l_1 and l_2 of Fieller's theorem. ☺

References

E. Fieller. The distribution of the index in a normal bivariate population. *Biometrika*, 24(3/4):428–440, 1932.

E. Fieller. Some problems in interval estimation. *J. Roy. Statist. Soc. Ser. B.*, 16:175–185, 1954.

V. H. Franz. Confidence limits for ratios. submitted 2004.

V. Guiard. Some remarks on the estimation of the ratio of the expectation values of a two-dimensional normal random variable (correction of the theorem of Milliken). *Biometrical J.*, 31(6):681–697, 1989.

D. V. Hinkley. On the ratio of two correlated normal random variables. *Biometrika*, 56(3):635 – 639, 1969.

D. V. Hinkley. Correction: On the ratio of two correlated normal random variables. *Biometrika*, 57(3):683, 1970.

M. Kendall and A. Stuart. *The Advanced Theory of Statistics*, volume 2. Charles Griffin & Company, London, 1961.

G. Marsaglia. Ratios of normal variables and ratios of sums of uniform variables. *Journal American Statistical Association*, 60(309):193 – 204, 1965.

G. Milliken. On a confidence interval about a parameter estimated by a ratio of normal random variables. *Comm. Statist. A—Theory Methods*, 11(17):1985–1995, 1982.

A. Rencher. *Multivariate Statistical Inference and Applications*. Wiley, New York, 1998.

M. Schervish. *Theory of Statistics*. Springer, New York, 1995.