"Sampling the reference set" revisited : in honour of Sir Ronald A. Fisher
van Berkum, E.E.M.; Linssen, H.N.; Overdijk, D.A.

Published: 01/01/1996

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Memorandum COSOR 96-26

"Sampling the reference set" revisited

in honour of Sir Ronald A. Fisher

E.E.M. van Berkum
H.N. Linssen
D.A. Overdijk

Eindhoven, August 1996
The Netherlands
"Sampling the reference set" revisited
in honour of Sir Ronald A. Fisher

Abstract

The confidence level of an inference table is defined as a weighted truth probability of the inference when sampling the reference set. The reference set is recognized by conditioning on the values of maximal partially ancillary statistics. In the sampling experiment values of incidental parameters are sampled from fiducial distributions. The concepts in this paper generalize ideas of Fisher (1961).

AMS Subject Classification 62A99, 62F99.

Keywords: Conditional inference, Inference table, Confidence level, Ancillarity, Fiducial distribution.
0 Introduction

What value to assign to the epistemological probability of statistical inference is still subject of debate. Some hold that it should be equated to the value of a posterior probability. Some may maintain that it is the truth probability of the inference when the same population is repeatedly sampled. Simple examples (Fisher (1973) and Dawid (1982)) show that the latter point of view may lead to unsound inferences. These examples serve to show that not the whole sample space but only a subset of it, the ‘reference set’, should be sampled, as a statistical experiment may be an ancillary mixture of subexperiments.

The reference set for any observation is contained in the sample space of the subexperiment to which the observation belongs. The recognition of the appropriate reference set is an important and necessary step in any process of verification of statistical inferences.

0.1 Sampling the reference set (Fisher, 1961)

Consider the inference concerning the mean of a normal population. The sample size \( n \) and the sample variance \( s^2 \) are known characteristics of the reference set. In the process of verification of any table for the population mean \( \mu \), the table is applied only if the generated sample has the same characteristics. Consider the table that gives the parameter statement \( \mu > \bar{x} - c.s/\sqrt{n} \) where \( c \) is some constant and \( \bar{x} \) represents the sample mean. Its conditional truth probability equals \( \Phi(c.s/\sigma) \) and depends on the incidental (or ‘nuisance’) parameter \( \sigma \).

Of course, some values of \( \sigma \) are more likely than others, given the observation \( s^2 \). Tables for \( \sigma \) can be defined by using the pivot \((n-1)s^2/\sigma^2\), where \( n \) is ancillary. These tables can be represented by a fiducial distribution for \( \sigma \), i.e. the confidence level of any statement for \( \sigma \) in any of the tables can be evaluated by integration of the fiducial density of \( \sigma \) over the region defined by the statement.

Now the process of verification can be defined. The observations can be sufficiently represented by \( n, \bar{x} \) and \( s^2 \). The reference set is sampled as follows. First an arbitrary value of \( \mu \) is chosen. Then a value of \( \sigma \) is sampled from its fiducial distribution. Now a value of \( \bar{x} \) is generated according to the simultaneous distribution of \( n, \bar{x} \) and \( s^2 \) with parameters \( \mu \) and \( \sigma^2 \), conditioned on the observed values of \( n \) and \( s^2 \). The table is applied, yielding a statement for \( \mu \) that may be either true or false. The confidence level of the table is now taken to be equal to the truth probability of this experiment.

The fiducial distribution for \( \sigma \) is verified similarly. Now \( n \) characterizes the reference set. An arbitrary value of \( \sigma \) is chosen and a value of \( s^2 \) is generated according to the simultaneous distribution of \( n \) and \( s^2 \), conditioned on the observed value of \( \mu \). The fiducial probability (or confidence level) of any statement for \( \sigma \) is the truth probability of this simple experiment.

0.2 Sampling the reference set (General)

The above well-known example is described here in some detail to illustrate the process of verification in an uncomplicated setting. In this article Fisher’s concept of sampling the reference set is generalized in order to be able to cope with more complicated settings. One complication is that the ancillary information (i.e. the observed value of the ancillary statistic and its distribution) may allow for inference with respect to only a part of the incidental parameter. Another is that the relevant truth probabilities of the table under consideration
may depend on only a part of the incidental parameter, with the logical consequence that inference is needed for that part only. The most interesting complication finally is that the ancillary subexperiment, to which the observation belongs may itself be an ancillary mixture of subexperiments and so on (Berkum et al. (1995)).

\[
\begin{align*}
\Omega & : \omega \\
\text{first level} & : A_1 \quad A_0 \\
\text{second level} & : A_{11} \quad A_{10}
\end{align*}
\]

The natural generalization of Fisher's verification experiment can now be described. Let us start with an example. Let \( \Omega \) be the set of possible outcomes of an experiment and let \( \omega \) be the actual outcome or observation. Suppose that the maximal partially ancillary statistic (see Section 2) divides the sample space into two subsets \( A_1 \) and \( A_0 \). This statistic is called ancillary on the first level. Suppose also that \( A_1 \) is divided into two subsets \( A_{11} \) and \( A_{10} \) by an ancillary statistic on the second level and let the observation \( \omega \) belong to \( A_{11} \) (see figure). With respect to the conditional model given \( A_{11} \) let there exist no nontrivial ancillary statistic. So \( A_{11} \) is the reference set corresponding to the observation \( \omega \). It is sampled as follows. First a value of the parameter is chosen, where the choice of the incidental parameter part should be compatible with the ancillary information on the first level. For example, in the Student case discussed by Fisher, the value of \( \mu \) is free to choose but the value of \( \sigma \) results from sampling its fiducial distribution. After the parameter value is determined an observation is generated from the corresponding conditional distribution given \( A_1 \). If this observation belongs to the reference set \( A_{11} \) the table under consideration is applied and the truth value of the corresponding parameter statement is determined. If it belongs to \( A_{10} \) it is discarded and a new value of the incidental parameter part is chosen. This time however the choice is made to be compatible with the ancillary information on the second level. Subsequently a new observation is sampled from the corresponding conditional distribution given the reference set \( A_{11} \). Then the table is applied.

So, in general, if a particular table should be verified a parameter value is chosen in a way that is compatible with the first level ancillary information. Sampling from the corresponding conditional distribution, given the appropriate ancillary subset, gives an observation that either belongs to the reference set or to an ancillary subset on a deeper level. In the latter case the current observation is discarded and the incidental parameter part is chosen in accordance with the ancillary information on the deepest possible level, whereas the interesting parameter part remains unchanged. A new observation is sampled from the appropriate conditional distribution. The process is repeated until the reference set is reached. If there are \( n \) ancillary levels the reference set is reached in at most \( n \) steps. If there is no nontrivial ancillary statistic then \( \Omega \) itself is the reference set. When the reference set is reached the table is applied and the truth value of the corresponding parameter statement is determined. The choice of the parameter value may influence the value of the truth probability associated with the table. If choices are done such that the truth probability is minimized, the resulting truth probability may be termed the confidence level of the table.

We are confident that the above description of the process of verification is generally applica-
ble. ‘Confidence’ is seen to be the value of the truth probability in a well-defined experiment. The original observation together with the specification of the interesting parameter part determines the subset of the sample space to be sampled. Which sampling distribution should be used depends on the ancillary information.

0.3 Supplementary remarks

- **Negative relevance**

  The ancillary structure of inference models is ill-conditioned in the sense that an infinitesimal change in the probability model may cause the ancillary structure to disappear. The examples of Fisher and Cox (Fisher (1973) and Dawid (1982)) can easily be modified such that this is the case. Unconditional inferences may now be unsound as negatively relevant subsets of the reference set appear. Approximate ancillary statistics may now be used to define suitable inferences.

- **Coherence**

  If we agree that a table is unsound when negatively relevant subsets of the reference set can be recognized, then statistical inferences, or rather the associated inferential probabilities, are not coherent. It is not possible to avoid positively relevant subsets for the two-sided Student inference table (see Brown (1967)) which implies that negatively relevant subsets can be recognized for the complimentary inference table.

  An even simpler instance of noncoherence of inferential probabilities occurs when one observation \( x \in \mathbb{R} \) of an unspecified distribution with median \( \theta \) is given. The inferential statement \( \theta \geq x \) has inferential probability \( \frac{1}{2} \), whereas the complimentary statement \( \theta < x \) has probability 0. Different inference rules for the same inference model may give rise to noncoherent inferential distributions, i.e. to a given inferential statement more than one inferential probability may be assigned, depending on the inference rule (see Wilkinson (1977) or Berkum et al. (1996)).

- **Invariance and sufficiency**

  The convention that observational evidence should be reduced by means of invariant or sufficient statistics has the effect that observations are grouped together and that the inference is the same for each member of the group. It excludes from consideration inferences that contain arbitrary elements contrary to common sense.

- **Statistical inference**

  The \( \sigma \)-field of interest is (see Section 3) generated by a collection of subsets of parameter space corresponding to a prespecified collection of admitted inferential statements. This collection is called the core of interest. Every inferential statement in an inference is an element of the core. The core could be all open convex sets for example.

  An inference table is called proper if it can be verified by appropriate sampling of the reference set and there do not exist negatively relevant subsets of the reference set.

  An inference table is called sharp (with respect to a given core) if there does not exist a
second proper table with the same level of confidence and such that the inferential state-
ment (an element of the core) for each observation is contained in the statement of the
first table.
It is tempting to state that an inference table is valid if and only if it is proper and sharp.

0.4 Summary

In Section 1 some measure theoretic definitions and theorems are given, that are to be used
in later sections. In Section 2 the maximal partially ancillary statistic is defined. The
ancillary trace and the reference set for any given observation in any given inference model
are determined in Section 3. In Section 4 the confidence level of a given inference rule is
defined by appropriate sampling the reference set. Section 5 contains some examples.

1 Preliminaries

In this section we collect preliminaries for reference in subsequent sections. Therefore the
reading of the paper can start with Section 2 and relevant parts of this section can be read
when they are referenced.

1.1 Let \( \lambda \) be a map from a set \( X \) into a set \( Y \). We refer to \( X \) as the domain and to \( Y \)
as the codomain of \( \lambda \). The powerset of \( X \), i.e. the collection of subsets of \( X \), is written as
\( P(X) \). The maps

\[
\lambda_\rightarrow : P(X) \rightarrow P(Y) ,
\lambda_\leftarrow : P(Y) \rightarrow P(X) ,
\]

are defined as

\[
\lambda_\rightarrow (A) := \{ \lambda(a) \in Y \mid a \in A \}, \quad A \subset X ,
\lambda_\leftarrow (B) := \{ b \in X \mid \lambda(b) \in B \}, \quad B \subset Y .
\]

Let \( \mathcal{F} \) be a \( \sigma \)-field of subsets of the codomain \( Y \), i.e. \( (Y, \mathcal{F}) \) is a measurable
space. We refer to

\[
(1.1.2) \quad \sigma(\lambda) := \{ \lambda_\leftarrow (B) \mid B \in \mathcal{F} \}
\]
as the \( \sigma \)-field of subsets of the domain \( X \) generated by the map \( \lambda \) from \( X \) into the measurable
space \( (Y, \mathcal{F}) \).

1.2 Let \( X \) be a set and let \( \mathcal{A} \subset P(X) \) be a collection of subsets of \( X \). The intersection of
the \( \sigma \)-fields on \( X \) containing \( \mathcal{A} \) is said to be the \( \sigma \)-field \( \sigma(\mathcal{A}) \) generated by \( \mathcal{A} \).
1.3 Let \((X, \mathcal{F})\) be a measurable space. For \(A \subseteq X\) we write

\[
(1.3.1) \quad \mathcal{F}|A := \{ B \cap A \mid B \in \mathcal{F} \},
\]

and we refer to the \(\sigma\)-field \(\mathcal{F}|A\) on \(A\) as the trace of \(\mathcal{F}\) on \(A\).

1.4 We now formulate a theorem for later use.

Theorem. Let \(\lambda\) be a surjection from a set \(X\) into the set \(X_1\) of a measurable space \((X_1, \mathcal{F}_1)\). The map \(\lambda^- : \mathcal{F}_1 \rightarrow \sigma(\lambda)\) is an isomorphism between the \(\sigma\)-fields \(\mathcal{F}_1\) on \(X_1\) and \(\sigma(\lambda)\) on \(X\), i.e. \(\lambda^-\) is a bijection from \(\mathcal{F}_1\) into \(\sigma(\lambda)\) such that

\[
\lambda^-(B \cap C) = \lambda^-(B) \cap \lambda^-(C)
\]

for all \(B, C \in \mathcal{F}_1\). Furthermore, the map \(\lambda^- : \sigma(\lambda) \rightarrow \mathcal{F}_1\) is the inverse map of the bijection \(\lambda^- : \mathcal{F}_1 \rightarrow \sigma(\lambda)\).

Proof. Obviously we have

\[
\lambda^-(B \cap C) = \lambda^-(B) \cap \lambda^-(C)
\]

for all \(B, C \in \mathcal{F}_1\). We now show that \(\lambda^- : \mathcal{F}_1 \rightarrow \sigma(\lambda)\) is a bijection. Let \(B \in \mathcal{F}_1\) and put \(A = \lambda^-(B) \in \sigma(\lambda)\). Evidently we have \(\lambda^-(A) \subseteq B\). Since \(\lambda : X \rightarrow X_1\) is a surjection we have \(\lambda^-(A) = B\). Hence \(\lambda^- : \mathcal{F}_1 \rightarrow \sigma(\lambda)\) is a bijection such that \(\lambda^- : \sigma(\lambda) \rightarrow \mathcal{F}_1\) is the inverse map. \(\square\)

Corollary. Let \((X_0, \mathcal{F}_0), (X_1, \mathcal{F}_1), (X_2, \mathcal{F}_2)\) be three measurable spaces and let \(\lambda\) be a measurable surjection from \((X_1, \mathcal{F}_1)\) into \((X_2, \mathcal{F}_2)\). The product \(\sigma\)-fields \(\mathcal{F}_0 \otimes \sigma(\lambda)\) and \(\mathcal{F}_0 \otimes \mathcal{F}_2\) are isomorphic. The corresponding isomorphism \(\tilde{\lambda}\) between the \(\sigma\)-field \(\mathcal{F}_0 \otimes \sigma(\lambda)\) on \(X_0 \times X_1\) and the \(\sigma\)-field \(\mathcal{F}_0 \otimes \mathcal{F}_2\) on \(X_0 \times X_2\) is given by

\[
\tilde{\lambda}(A) = \{(x_0, \lambda(x_1)) \mid (x_0, x_1) \in A\} \in \mathcal{F}_0 \otimes \mathcal{F}_2, \quad A \in \mathcal{F}_0 \otimes \sigma(\lambda).
\]

1.5 Let \(X\) be a set and let \(I\) be an index set such that to every \(i \in I\) there correspond a measurable space \((X_i, \mathcal{F}_i)\) and a map \(\lambda_i : X \rightarrow X_i\). The \(\sigma\)-field on \(X\) generated by the collection

\[
\Lambda := \{ \lambda_i \mid i \in I \}
\]

of maps on \(X\) is written as \(\sigma(\Lambda)\) and

\[
(1.5.1) \quad \sigma(\Lambda) := \sigma(\bigcup_{i \in I} \sigma(\lambda_i));
\]

see 1.2 and (1.1.2).
Let \((X, \mathcal{F})\) be a measurable space. The set of probability measures on \(\mathcal{F}\) is denoted by \(\mathcal{F}'\). For \(A \in \mathcal{F}\) consider the map

\[(1.6.1) \quad \lambda_A(p) := p(A) \in [0,1], \quad p \in \mathcal{F}'.\]

from \(\mathcal{F}'\) into the unit interval. The \(\sigma\)-field of Borel subsets of the unit interval is written as \(\mathcal{B}[0,1]\). The \(\sigma\)-field generated by the collection

\[\{\lambda_A \mid A \in \mathcal{F}\}\]

denotes the \(\sigma\)-field of Borel subsets of the measurable space \([0,1],[\mathcal{B}[0,1]]\) and is denoted by \(\mathcal{F}(\mathcal{F})\); see (1.5.1). We refer to \((\tilde{\mathcal{F}}, \mathcal{F})\) as the space of probability measures on the \(\sigma\)-field \(\mathcal{F}\). Let \(\mathcal{P} \subseteq \tilde{\mathcal{F}}\) be a collection of probability measures on \(\mathcal{F}\). We refer to \((\mathcal{P}, \mathcal{F}|\mathcal{P})\) as a space of probability measures on \(\mathcal{F}\); see (1.3.1).

Let \(\mathcal{F}_1, \mathcal{F}_2\) be two \(\sigma\)-fields on a set \(X\). The \(\sigma\)-fields \(\mathcal{F}_1\) and \(\mathcal{F}_2\) are called algebraically independent, notation \(\mathcal{F}_1 \perp \mathcal{F}_2\), if every inclusion between sets of \(\mathcal{F}_1\) and \(\mathcal{F}_2\) is trivial, i.e. for all \(A \in \mathcal{F}_1\) and \(B \in \mathcal{F}_2\) we have

\[(1.7.1) \quad A \subseteq B \Rightarrow A = \emptyset \text{ or } B = X.\]

Let \((X, \mathcal{F})\) be a measurable space and let \(\mathcal{P} \subseteq \tilde{\mathcal{F}}\) be a family of probability measures on \(\mathcal{F}\). The space of all probability measures on \(\mathcal{F}\) is written as \((\tilde{\mathcal{F}}, \mathcal{F})\); see (1.6). The \(\sigma\)-field \(\mathcal{F}_0 \subseteq \mathcal{F}\) on \(X\) is said to be regular, if there exists a measurable map \(P\) from \((X \times \mathcal{P}, \mathcal{F}_0 \otimes \mathcal{F}|\mathcal{P})\) into \((\tilde{\mathcal{F}}, \mathcal{F})\) such that

\[(1.8.1) \quad \int_A P(x,p)(B)p(dx) = p(A \cap B)\]

for all \(A \in \mathcal{F}_0, B \in \mathcal{F}\) and \(p \in \mathcal{P}\). Here \(\mathcal{F}_0 \otimes \mathcal{F}|\mathcal{P}\) is the product \(\sigma\)-field on the cartesian product \(X \times \mathcal{P}\) corresponding to the \(\sigma\)-fields \(\mathcal{F}_0\) and \(\mathcal{F}|\mathcal{P}\) on \(X\) and \(\mathcal{P}\) respectively. For \(p \in \mathcal{P}\) we refer to the map

\[(1.8.2) \quad P(\cdot, p): (X, \mathcal{F}_0) \to (\tilde{\mathcal{F}}, \mathcal{F})\]

as the regular conditional probability measure on \(\mathcal{F}\) given \(\mathcal{F}_0\) corresponding to \(p \in \mathcal{P}\).

A measurable map \(S\) from \((X, \mathcal{F})\) into a measurable space is said to be a statistic on \((X, \mathcal{F})\), if the \(\sigma\)-field \(\sigma(S) \subseteq \mathcal{F}\) on \(X\) generated by \(S\) is regular.

Let \((X, \mathcal{F})\) be a measurable space. The nonempty set \(A \in \mathcal{F}\) is said to be an atom in \(\mathcal{F}\), if for all \(B \in \mathcal{F}\) we have

\[(1.9.1) \quad B \subseteq A \Rightarrow B = \emptyset \text{ or } B = A.\]
The σ-field $\mathcal{F}$ on $X$ is called atomic, if there exists a partition of $X$ such that every element of the partition is an atom in $\mathcal{F}$.

Let $(X, \mathcal{F})$ be a measurable space. The property that $\mathcal{F}$ is atomic does not imply that a sub-σ-field $\mathcal{F}_0 \subset \mathcal{F}$ is atomic also. This can be seen in the following example.

Consider the measurable space $(\mathbb{R}, \mathcal{F})$, where the σ-field $\mathcal{F}$ on the set $\mathbb{R}$ of the real numbers is given by

$$\mathcal{F} := \{A \subset \mathbb{R} \mid A \text{ is countable or } \mathbb{R} \setminus A \text{ is countable}\}.$$

Obviously the σ-field $\mathcal{F}$ on $\mathbb{R}$ is atomic. Now consider

$$\mathcal{F}_0 := \{A \in \mathcal{F} \mid A \cap [0, 1] = \emptyset \text{ or } A \subset [0, 1]\}.$$

It is easily verified that $\mathcal{F}_0 \subset \mathcal{F}$ is a nonatomic σ-field on $\mathbb{R}$.

2 Partially ancillary statistic

In this section the relevant terminology is introduced and the definition of a maximal partially ancillary statistic is given.

Consider an experiment. The set of possible outcomes of the experiment is denoted by $\Omega$. The σ-field of events on $\Omega$ is written as $\Sigma$. The measurable space $(\Omega, \Sigma)$ is the sample space of the experiment. Let $(\hat{\Sigma}, \widetilde{\Sigma})$ be the space of probability measures on $\Sigma$; see 1.6. The probability distribution on $\Sigma$ corresponding to the outcome of the experiment is not known. However, a subset $\mathcal{P} \subset \hat{\Sigma}$ is given such that the probability distribution of the outcome of the experiment is in the set $\mathcal{P}$. The measurable space $(\mathcal{P}, \widetilde{\Sigma}|\mathcal{P})$, see 1.6, of probability measures on $\Sigma$ is said to be the probability model of the experiment. We assume that every sufficient and every ancillary statistic is trivial. If this is not the case, then the observation should be reduced to the value of the minimal sufficient statistic and its distribution should be conditioned on the value of a maximal ancillary statistic. Let $p_0$ be the probability distribution of the outcome of the experiment. It may be that one is interested only in a specific aspect of $p_0$, i.e. a collection of admitted inferential statements is specified. The σ-field generated by this collection is referred to as the σ-field of interest $\mathcal{R} \subset \widetilde{\Sigma}|\mathcal{P}$. The triple

$$\begin{align*}
\text{sample space} & \quad (\Omega, \Sigma), \\
\text{probability model} & \quad (\mathcal{P}, \widetilde{\Sigma}|\mathcal{P}), \\
\text{σ-field of interest} & \quad \mathcal{R} \subset \widetilde{\Sigma}|\mathcal{P}
\end{align*}$$

is said to constitute an inference model for the experiment. The authors introduced a reduction and conditioning scheme, the so-called SCIRA scheme, to transform inference models into so-called reference models (Berkum et al. (1995)). In a reference model all invariant, partially sufficient and partially ancillary statistics are trivial. In the process of verification of statistical methods as described in this paper, the conditioning steps (A-steps) in SCIRA play a crucial role. Therefore in this section we recall the definition of a partially ancillary
Let $A$ be a statistic on the sample space $(\Omega, \Sigma)$. This statistic induces a partition of $\Omega$. Its elements are given by

$$ (2.2) \quad C(\omega) := \{ \eta \in \Omega \mid A(\eta) = A(\omega) \}, \quad \omega \in \Omega. $$

The $\sigma$-field on $\Omega$ generated by $A$ is denoted by $\sigma(A)$; see (1.1.2). We suppose that $\sigma(A)$ is atomic, such that $C(\omega)$ is an atom of $\sigma(A)$ for all $\omega \in \Omega$; see 1.9. The marginal probability distribution on $\sigma(A)$ corresponding to $\varphi(\omega)$ is denoted by $\varphi(p)$. The $\sigma$-field on $\mathcal{P}$ generated by the map $\varphi$ from $\mathcal{P}$ into the space $(\sigma(A), \sigma(A))$ of probability measures on $\sigma(A)$, see 1.6, is written as $\sigma(\varphi) \subseteq \Sigma|\mathcal{P}$.

Fix $\omega \in \Omega$. For $p \in \mathcal{P}$ we write

$$ (2.3) \quad \psi(p) := P(\omega, p) \in \widehat{\Sigma}, $$

where $P(\cdot, p) : (\Omega, \sigma(A)) \rightarrow (\widehat{\Sigma}, \overline{\Sigma})$ is a regular conditional probability distribution on $\Sigma$ given $\sigma(A)$ corresponding to $p \in \mathcal{P}$; see (1.8.2). The $\sigma$-field on $\mathcal{P}$ generated by the map $\psi$ from $\mathcal{P}$ into the space $(\widehat{\Sigma}, \overline{\Sigma})$ of probability measures on $\Sigma$ is denoted by $\sigma(\psi) \subseteq \overline{\Sigma}|\mathcal{P}$.

The statistic $A$ on $(\Omega, \Sigma)$ is said to be partially ancillary with respect to $\omega \in \Omega$, if the marginal distribution $\varphi(p)$ is not informative with respect to the interesting aspect of $p \in \mathcal{P}$, i.e.

$$ (2.4) \quad \mathcal{R} \perp \sigma(\varphi); $$

see 1.7, and the interesting aspect of $p \in \mathcal{P}$ is a function of the conditional probability distribution $\psi(p)$, i.e. the map $\psi$ in (2.3) can be chosen such that

$$ (2.5) \quad \mathcal{R} \subseteq \sigma(\psi). $$

The partially ancillary statistic $A$ is said to be maximal, if for all partially ancillary statistics $A'$ on $(\Omega, \Sigma)$ we have

$$ (2.6) \quad \sigma(A) \subseteq \sigma(A') \Rightarrow \sigma(A) = \sigma(A'). $$

In general there does not exist a unique maximal partially ancillary statistic with respect to $\omega \in \Omega$.

## 3 Ancillary trace and reference set

The first important step in the verification of inference tables is the determination of the reference set. A statistical experiment may be an ancillary mixture of subexperiments. The mixing distribution is not informative with respect to the interesting parameter part, but may be informative with respect to the incidental parameter part. By use of a maximal partially ancillary statistic, as defined in Section 2, the subexperiment to which the original observation belongs is determined. The subexperiment may again be an ancillary mixture.
of subexperiments and so on. At every step sample space is reduced by means of invariant or partially sufficient statistics. In this way we construct a decreasing sequence of subsets of sample space, such that every element of the sequence can be considered as the sample space of a subexperiment. This chain of subsets is called an ancillary trace. When the subexperiment is not a mixture, the reference set is recognized to be the smallest element of the chain.

Let \((\Omega, \Sigma)\) be the sample space, \((\mathcal{P}, \Sigma|\mathcal{P})\) the probability model and \(\mathcal{R} \subset \Sigma|\mathcal{P}\) the \(\sigma\)-field of interest of an inference model; see (2.1). Furthermore, let \((\Omega, \Sigma_0), (\mathcal{P}_0, \Sigma_0|\mathcal{P}_0), \mathcal{R}_0 \subset \Sigma_0|\mathcal{P}_0\) be the inference model in the SCIRA reduction and conditioning scheme before the first conditioning step on a maximal partially ancillary statistic. So the \(\sigma\)-field \(\Sigma_0 \subset \Sigma\) corresponds to invariant and partially sufficient reduction steps in SCIRA. The set \(\mathcal{P}_0\) contains the marginal probability distributions on \(\Sigma_0\) corresponding to the probability distributions on \(\Sigma\) contained in \(\mathcal{P}\), i.e.

\[
(3.1) \quad \mathcal{P}_0 = \{\psi_0(p) \in \tilde{\Sigma}_0 \mid p \in \mathcal{P}\}.
\]

where \(\psi_0(p)\) is the marginal probability distribution on \(\Sigma_0\) corresponding to \(p \in \mathcal{P}\). For the \(\sigma\)-field \(\mathcal{R}_0\) of interest we have

\[
(3.2) \quad \mathcal{R}_0 = \{\psi_0^{-1}(B) \in \tilde{\Sigma}_0|\mathcal{P}_0 \mid B \in \mathcal{R}\};
\]

see 1.1. The \(\sigma\)-fields \(\mathcal{R}\) and \(\mathcal{R}_0\) are isomorphic (see Berkum et al. (1995)). Fix the observation \(\omega \in \Omega\) and define

\[
(3.3) \quad C_0(\omega) := \Omega.
\]

The SCIRA reduction and conditioning scheme of the inference model \((\Omega, \Sigma_0), (\mathcal{P}_0, \Sigma_0|\mathcal{P}_0), \mathcal{R}_0 \subset \Sigma_0|\mathcal{P}_0\) is the repeated application of a transformation. Essentially, this basic transformation is performed in two steps. The first step is conditioning on a nontrivial maximal partially ancillary statistic with respect to \(\omega\), the second step is a reduction corresponding to invariant and/or partially sufficient statistics. Suppose that after \(k, k \geq 0\), repetitions of the above transformation, we obtain a reference model with respect to \(\omega\), i.e. all invariant, partially sufficient and partially ancillary statistics are trivial. So \(k + 1\) inference models are specified. We now recursively introduce notations related to these inference models. Let

\[
(\Omega, \Sigma_{i-1}), (\mathcal{P}_{i-1}, \Sigma_{i-1}|\mathcal{P}_{i-1}), \mathcal{R}_{i-1} \subset \Sigma_{i-1}|\mathcal{P}_{i-1}
\]

be the inference model after \(i - 1, i = 1, 2, \ldots, k\), transformations, and let \(A_i\) be the nontrivial maximal partially ancillary statistic on \((\Omega, \Sigma_{i-1})\) corresponding to the conditioning step in the \(i^{th}\)-transformation. Furthermore, let \(\Sigma_i \subset \Sigma_{i-1}\) be the \(\sigma\)-field on \(\Omega\) corresponding to the reduction step in the \(i^{th}\)-transformation. So the sample space of the new inference model is \((\Omega, \Sigma_i)\). Since in (3.3) we defined \(C_0(\omega) := \Omega\), we may assume to have constructed recursively the sets

\[
\Omega = C_0(\omega) \supset C_1(\omega) \supset \cdots \supset C_i(\omega).
\]
Define

\[(3.4) \quad C_i(\omega) := \{ \eta \in C_{i-1}(\omega) \mid A_i(\eta) = A_i(\omega) \} \in \sigma(A_i), \]

where \( \sigma(A_i) \subset \Sigma_{i-1} \) is the \( \sigma \)-field on \( \Omega \) generated by \( A_i \). Since \( \Sigma_{i-1} \subset \Sigma \) the statistic \( A_i \) on \((\Omega, \Sigma_{i-1})\) may be considered as a statistic on \((\Omega, \Sigma)\) also. For \( p \in \mathcal{P} \) let

\[ P(\cdot, p) : (\Omega, \sigma(A_i)) \to (\bar{\Sigma}, \bar{\Sigma}) \]

be a regular conditional probability distribution on \( \Sigma \) given \( \sigma(A_i) \) corresponding to \( p \in \mathcal{P} \); see (1.8.2). For \( p \in \mathcal{P} \) we define the probability measure \( \psi_i(p) \) on \( \Sigma_i \) as follows

\[(3.5) \quad \psi_i(p)(B) := P(\omega, p)(B), \quad B \in \Sigma_i. \]

So we defined the map

\[ \psi_i : (\mathcal{P}, \bar{\Sigma}|\mathcal{P}) \to (\bar{\Sigma}_i, \bar{\Sigma}_i), \]

and for all \( p \in \mathcal{P} \) we have

\[(3.6) \quad \psi_i(p)(C_i(\omega)) = 1. \]

The probability model of the new inference model is written as \((\mathcal{P}_i, \bar{\Sigma}_i|\mathcal{P}_i)\), where

\[(3.7) \quad \mathcal{P}_i := \{ \psi_i(p) \in \bar{\Sigma}_i \mid p \in \mathcal{P} \}. \]

and the \( \sigma \)-field \( \mathcal{R}_i \subset \bar{\Sigma}_i|\mathcal{P}_i \) of interest in the new inference model can be put in the form

\[(3.8) \quad \mathcal{R}_i := \{ \psi_i^{-1}(B) \in \bar{\Sigma}_i|\mathcal{P}_i \mid B \in \mathcal{R} \}. \]

The map

\[ \psi_i : (\mathcal{P}, \bar{\Sigma}|\mathcal{P}) \to (\mathcal{P}_i, \bar{\Sigma}_i|\mathcal{P}_i) \]

is a surjection and \( \mathcal{R} \subset \sigma(\psi_i) \), see (2.5). So it follows from the theorem in 1.4 that the \( \sigma \)-fields \( \mathcal{R} \) and \( \mathcal{R}_i \) are isomorphic. We now have constructed the inference model

\[(\Omega, \Sigma_i), \quad (\mathcal{P}_i, \bar{\Sigma}_i|\mathcal{P}_i), \quad \mathcal{R}_i \subset \bar{\Sigma}_i|\mathcal{P}_i, \quad 0 \leq i \leq k. \]

such that
\[ \Sigma_k \subset \Sigma_{k-1} \subset \cdots \subset \Sigma_1 \subset \Sigma_0 \subset \Sigma , \]

\[(3.10) \quad C_k(\omega) \subset C_{k-1}(\omega) \subset \cdots \subset C_1(\omega) \subset C_0(\omega) = \Omega ,
\]

\[ p(C_i(\omega)) = 1, \quad p \in \mathcal{P}_i, \quad 0 \leq i \leq k , \]

and the \( \sigma \)-fields \( \mathcal{R}_i \) and \( \mathcal{R} \), \( 0 \leq i \leq k \), are isomorphic. We refer to (3.10) as an ancillary trace with respect to the observation \( \omega \in \Omega \). The number \( k \) is called the trace depth. The inference model

\[
(\Omega, \Sigma_k), \quad (\mathcal{P}_k, \Sigma_k|\mathcal{P}_k), \quad \mathcal{R}_k \subset \Sigma_k|\mathcal{P}_k
\]

is a reference model with respect to \( \omega \) and the reference set \( \text{Ref}(\omega) \) is defined by

\[(3.11) \quad \text{Ref}(\omega) := C_k(\omega) .
\]

### 4 Sampling the reference set

The confidence level of a given inference table is defined as a weighted truth probability of the inference when sampling the reference set. These truth probabilities may depend on the incidental parameters. This dependency determines the interesting aspect of the incidental parameter. In this context the concept of crosscoherence plays a crucial role. The incidental problem is an inference problem as well. This allows for a recursive definition of confidence level. We first describe the ancillary structure of the inference problem.

Let \( (\Omega, \Sigma) \) be the sample space, \( (\mathcal{P}, \Sigma|\mathcal{P}) \) the probability model and \( \mathcal{R} \subset \Sigma|\mathcal{P} \) the \( \sigma \)-field of interest of an inference model. We suppose that the \( \sigma \)-field \( \mathcal{R} \) of interest is atomic, i.e. there exists a partition

\[(4.1) \quad \mathcal{P} = \bigcup_{\gamma \in \Gamma} R_\gamma
\]

of the set \( \mathcal{P} \) such that for every \( \gamma \) in the index set \( \Gamma \) the set \( R_\gamma \in \mathcal{R} \) is an atom in the \( \sigma \)-field \( \mathcal{R} \) of interest; see 1.9.

For \( \omega \in \Omega \) let

\[(4.2) \quad (\Omega, \Sigma_i), \quad (\mathcal{P}_i, \Sigma_i|\mathcal{P}_i), \quad \mathcal{R}_i \subset \Sigma_i|\mathcal{P}_i , \quad 0 \leq i \leq k(\omega) ,
\]

be the \( k(\omega) + 1 \) inference models corresponding to the ancillary trace

\[(4.3) \quad \text{Ref}(\omega) = C_{k(\omega)}(\omega) \subset C_{k(\omega)-1}(\omega) \subset \cdots \subset C_1(\omega) \subset C_0(\omega) = \Omega ;
\]

see (3.9-11). The marginal probability distribution on \( \Sigma_0 \) corresponding to \( p \in \mathcal{P} \) is denoted by \( \psi_0(p) \in \mathcal{P}_0 \); see (3.1). For \( 1 \leq i \leq k(\omega) \) the conditional probability distribution on \( \Sigma_i \) given the value \( A_i(\omega) \) of the partially ancillary statistic \( A_i \) on \( (\Omega, \Sigma_{i-1}) \) corresponding to \( p \in \mathcal{P} \) is written as \( \psi_i(p) \in \mathcal{P}_i \); see (3.7). So for \( 0 \leq i \leq k(\omega) \)
4.1 Inference table

A table fixes a subset \( U^\omega \) of the parameter space \( \mathcal{P} \) for every observation \( \omega \in \Omega \), such that \( U^\omega \in \mathcal{R} \). To the observation \( \omega \in \Omega \) corresponds the inferential statement \( p_0 \in U^\omega \), where \( p_0 \in \mathcal{P} \) is the unknown probability distribution of the outcome of the experiment. So a table is represented as a set \( U \subset \Omega \times \mathcal{P} \) such that \( U \in \Sigma \otimes \mathcal{R} \), where \( \Sigma \otimes \mathcal{R} \) is the product \( \sigma \)-field on the cartesian product \( \Omega \times \mathcal{P} \) corresponding to the \( \sigma \)-fields \( \Sigma \) and \( \mathcal{R} \) on \( \Omega \) and \( \mathcal{P} \) respectively. For \((\omega, p) \in \Omega \times \mathcal{P} \) we have

\[
(4.1.1) \quad U^\omega = \{ p \in \mathcal{P} \mid (\omega, p) \in U \} \in \mathcal{R} ,
\]

and we write

\[
(4.1.2) \quad U_p := \{ \omega \in \Omega \mid (\omega, p) \in U \} \in \Sigma .
\]

So \( U_p \) is the set of observations \( \omega \) such that \( p \in U^\omega \). Suppose that two observations \( \omega_1, \omega_2 \in \text{Ref}(\omega) \), \( \omega \in \Omega \), must be identified according to the reduction of \( \Sigma \) to \( \Sigma_k(\omega) \). A table \( U \) is assumed to satisfy

\[
(4.1.3) \quad U \in \Sigma_k(\omega) \otimes \mathcal{R} .
\]

4.2 Confidence level

In the verification experiment for a table \( U \) as described in Section 0.2 the value of the interesting parameter part remains unchanged. So for a chosen \( \gamma \in \Gamma \) only the probability distributions in the atom \( R_\gamma \subset \mathcal{P} \) of the \( \sigma \)-field \( \mathcal{R} \) of interest are used as sampling distributions in the verification experiment; see (4.1).

Let \( \omega \in \Omega \) be the original observation and let \( \omega^* \in \text{Ref}(\omega) \) be the outcome of the sampling experiment as described in 0.2 and to be specified in this section. Obviously we have \( U^\omega \in \mathcal{R} \); see (4.1.1). So for the atom \( R_\gamma \) in \( \mathcal{R} \) either \( R_\gamma \subset U^\omega \) or \( R_\gamma \subset \mathcal{P} \setminus U^\omega \). We propose the confidence level associated with the table \( U \) and the original observation \( \omega \in \Omega \) to be essentially the truth probability of the statement

\[
(4.2.1) \quad R_\gamma \subset U^\omega .
\]

We refer to this statement as the sampled inference corresponding to \( \gamma \in \Gamma \). We now specify the sampling experiment.
In the verification experiment we consider \( k(\omega) + 1 \) subexperiments. The inference model

\[
(\Omega, \Sigma_i, (P_i, \Sigma_i|P_i), \mathcal{R}_i \subset \Sigma_i|P_i)
\]

corresponds to the \( i \)th subexperiment; see (4.2). The table to be verified in the \( i \)th subexperiment is given by

\[
U(i) := \{(\omega, \psi_i(p)) \in \Omega \times P_i \mid (\omega, p) \in U\};
\]

see (4.4). Note that \( U \in \Sigma_{k(\omega)} \otimes \mathcal{R}, \) \( U(i) \in \Sigma_{k(\omega)} \otimes \mathcal{R}_i \) and that the \( \sigma \)-fields \( \Sigma_{k(\omega)} \otimes \mathcal{R} \) and \( \Sigma_{k(\omega)} \otimes \mathcal{R}_i \) are isomorphic such that \( U \) and \( U(i) \) are each others images; see the corollary in 1.4. Let \( \alpha_i(\gamma, U, \omega) \) be essentially the truth probability of the sampled inference in the \( i \)th subexperiment for the table \( U(i) \) corresponding to \( \gamma \in \Gamma \) and the original observation \( \omega \in \Omega \). The 0th subexperiment is the verification experiment itself, and therefore the confidence level \( \alpha(U, \omega) \) associated with the table \( U \) and the original observation \( \omega \in \Omega \) is defined by

\[
(4.2.3) \quad \alpha(U, \omega) = \inf_{\gamma \in \Gamma} \alpha_0(\gamma, U, \omega).
\]

In the sequel the values of the truth probabilities \( \alpha_i, \) \( 0 \leq i \leq k(\omega), \) are calculated.

### 4.2.1 Truth probabilities

The values of the truth probabilities \( \alpha_i, \) \( 0 \leq i \leq k(\omega), \) are calculated by use of induction with respect to the trace depth \( k(\omega). \) First consider the case \( k(\omega) = 0. \) So every partially ancillary statistic is trivial. Now the reference set \( \text{Ref}(\omega) = \Omega \) and it is sampled by use of a probability distribution arbitrarily chosen in \( \mathcal{R}_i \subset P. \) Hence

\[
(4.2.1.1) \quad \alpha_0(\gamma, U, \omega) = \inf_{p \in \mathcal{R}_i} p(U_p);
\]

see (4.1.2). The reader who is not inclined to equate the confidence level of the table \( U \) to

\[
(4.2.1.2) \quad \alpha(U, \omega) = \inf_{\gamma \in \Gamma} \alpha_0(\gamma, U, \omega) = \inf_{p \in \mathcal{P}} p(U_p),
\]

may not benefit from further reading of the paper. Note that, in case \( k(\omega) = 0, \) the confidence level of \( U \) is independent of \( \omega \) and that it is not equal to the inference's repeated sampling truth probability itself but to its largest lower bound.

Assume that we have calculated the values of the truth probabilities \( \alpha_i, \) \( 0 \leq i \leq k(\omega), \) for all experiments with trace depth \( k(\omega) = n \) for fixed \( n \geq 0. \) We now calculate these truth probabilities for experiments with trace depth \( n + 1. \) According to the above assumption the values of the truth probabilities \( \alpha_i(U, \omega), \) \( 1 \leq i \leq k(\omega) = n + 1, \) are known. The value of \( \alpha_0 \) is computed in the following way. Let \( p \in \mathcal{R}_i \) be the sampling distribution of a verification experiment to be used in the first subexperiment. Fisher proposed to sample from the conditional distribution \( \psi_1(p) \in \mathcal{P}_1 \) given \( C_1(\omega); \) see (4.3). The sampled observation either
belongs to the reference set $\text{Ref}(\omega) = C_{n+1}(\omega)$ or to one of the sets $C_i(\omega) \setminus C_{i+1}(\omega)$, $1 \leq i \leq n$. The corresponding conditional truth probabilities are equal to $\psi_1(p)(U_p \mid \text{Ref}(\omega))$ and to $\alpha_i(\gamma, U, \omega)$ respectively; see (4.1.2) and (4.2.1). So for $p \in R_\gamma$ the truth probability is equal to

$$\beta(p) = \psi_1(p)(\text{Ref}(\omega)) \psi_1(p)(U_p \mid \text{Ref}(\omega)) + \sum_{i=1}^{n} \psi_1(p)(C_i(\omega) \setminus C_{i+1}(\omega)) \alpha_i(\gamma, U, \omega).$$

The incidental aspect of $p$ is its variation within $R_\gamma$. It may affect the value of the truth probability $\beta$. The ancillary statistic’s value and distribution provide an inferential distribution for this effect. For fixed $\gamma$, Fisher proposed to sample the incidental aspect of $p$ from this distribution and to sample an observation from the resulting distribution. The truth probability of this compound experiment is a weighted average of $\beta$. In general, the weighting distribution for $\beta$ is not unique. Therefore $\alpha_0(\gamma, U, \omega)$ is equal to the largest lower bound of the corresponding truth probabilities. To construct this weighting distribution it is necessary first to define the so-called incidental experiment.

4.2.1.1. Incidental experiment

The $\sigma$-field generated by the partially ancillary statistic $A_1$ on $(\Omega, \Sigma_0)$ is denoted by $\sigma(A_1) \subset \Sigma_0$. The sample space of the incidental experiment is $(\Omega, \sigma(A_1))$. The marginal probability distribution on $\sigma(A_1)$ corresponding to $p \in P_0$ is denoted by $\varphi(p)$. The $\sigma$-field on $P_0$ generated by the map $\varphi$ from $P_0$ into the space $(\sigma(A_1), \sigma(A_1))$ of probability measures on $\sigma(A_1)$ is denoted by $\sigma(\varphi) \subset \Sigma_0|P_0$. We have

$$\sigma(\varphi) \perp R_0 ;$$

see (2.4). Introduce the map

$$\lambda : (P, \Sigma|P) \rightarrow (\sigma(A_1), \sigma(A_1))$$

by

$$\lambda(p) := \varphi(p), \quad p \in P.$$ 

where $\varphi(p)$ is the marginal distribution on $\Sigma_0$ corresponding to $p \in P$; see (3.1). The $\sigma$-field on $P$ generated by the map $\lambda$ from $P$ into $(\sigma(A_1), \sigma(A_1))$ is written as $\sigma(\lambda) \subset \Sigma|P$. The $\sigma$-fields $R \subset \Sigma|P$ and $R_0 \subset \Sigma_0|P_0$ are isomorphic and therefore we conclude from (4.2.1.1.1)

$$\sigma(\lambda) \perp R .$$

The probability model of the incidental experiment can be written as

$$\lambda = (J, \sigma(A_1)|J),$$

where
\[ J := \{ \lambda(p) \in \sigma(A_1) \mid p \in \mathcal{R}_\gamma \} . \]

The collection \( J \) is independent of \( \gamma \), since it follows from \( \sigma(\lambda) \perp \mathcal{R} \) that

\[ \lambda^{-1}(\lambda(p)) \cap R_\delta \neq \emptyset \]

for all \( \delta \in \Gamma \) and \( p \in \mathcal{P} \).

The probability structure of the incidental experiment is given by

- sample space: \( (\Omega, \sigma(A_1)) \),
- probability model: \( (J, \sigma(A_1)|J) \).

We assume that every sufficient and every ancillary statistic on \( (\Omega, \sigma(A_1)) \) is trivial. If this is not the case, then \( \sigma(A_1) \) should be reduced according to the minimal sufficient statistic and the distribution of the observation in the incidental experiment should be conditioned on the value of a maximal ancillary statistic. We now establish the \( \sigma \)-field of interest in the incidental experiment.

4.2.1.1. Interest specification and crosscoherence

For the specification of the interest in the incidental experiment the \( \sigma \)-fields \( \sigma(\lambda)|R_\gamma \) and \( \sigma(\beta) \) on \( R_\gamma \) are relevant. Here \( \sigma(\beta) \) is the \( \sigma \)-field on \( R_\gamma \) generated by the map \( \beta \) in (4.2.1.3) from \( R_\gamma \) into the unit interval. The incidental experiment is introduced to construct the appropriate weighting distribution for the function \( \beta \) on \( R_\gamma \). Therefore the \( \sigma \)-field \( \sigma(\lambda)|R_\gamma \) should be reduced in relation to \( \sigma(\beta) \). For this reduction we use the concept of crosscoherence; compare Wilkinson (1977).

Two components of the \( \sigma \)-field \( \sigma(\lambda)|R_\gamma \) can be distinguished: one is relevant for \( \sigma(\beta) \) whereas the other is not. Let the \( \sigma \)-fields \( F_0 \) and \( F_1 \) be such that

\[ \sigma(\lambda)|R_\gamma = \sigma(F_0 \cup F_1) . \]

and the \( \sigma \)-field \( F_1 \) is not informative with respect to \( \sigma(\sigma(\beta) \cup F_0) \), i.e.

\[ F_1 \perp \sigma(\sigma(\beta) \cup F_0) ; \]

see 1.7. The latter condition cannot be replaced by the weaker condition \( F_1 \perp \sigma(\beta) \) and \( F_1 \perp F_0 \); see Example 5.5. Now the sub-\( \sigma \)-field \( F_0 \) of \( \sigma(\lambda)|R_\gamma \) is said to be crosscoherent with \( \sigma(\beta) \) if \( F_0 \) is minimal with respect to inclusion, i.e. for every reduction \( F_0^* \) of \( \sigma(\lambda)|R_\gamma \) in the above sense we have

\[ F_0^* \subset F_0 \Rightarrow F_0^* = F_0 . \]

We now specify the \( \sigma \)-field \( I_\gamma \) of interest in the incidental experiment. Let \( F_\gamma \subset \sigma(\lambda)|R_\gamma \) be crosscoherent with \( \sigma(\beta) \). The \( \sigma \)-field \( F_\gamma \) identifies the interest in the incidental experiment. It follows from the theorem in 1.4 that the \( \sigma \)-fields \( F_\gamma \) and
are isomorphic. So \( I_\gamma \) is the \( \sigma \)-field of interest in the incidental experiment, and the inference model in the incidental experiment is given by

- **Sample space** \( (\Omega, \sigma(A_1)) \)
- **Probability model** \( (J, \sigma(A_1)|J) \)
- **a-field of interest** \( I_\gamma \subseteq \sigma(A_1)|J \)

**Remark.** Failure to take crosscoherence into account resulted in the unsharp inference rule for the Behrens-Fisher problem (Fisher (1961)) as expressed in the New Cambridge Elementary Statistical Tables (1984). The sharper inference (Welch (1947)) to be found in the Biometrika Tables for Statisticians (1970) is not proper. Linssen (1991) claims to give an inference rule for the Behrens-Fisher problem that is both sharp and proper (see 0.3).

### 4.2.1.1.2 Fiducial averaging

The appropriate weighting distribution for \( \beta \) on \( R_\gamma \) is constructed by use of inferential distributions. These inferential distributions are discussed in Berkum et al. (1996). The incidental experiment is an inference problem itself. So the SCIRA reduction and conditioning scheme should be applied. First we discuss the case that the trace depth equals zero for every element of the sample space of the incidental experiment.

According to Berkum et al. (1996) we introduce the reference space

\[
(\Omega \times J, \sigma(A_1) \otimes I_\gamma)
\]

where \( \sigma(A_1) \otimes I_\gamma \) is the product \( \sigma \)-field on the cartesian product \( \Omega \times J \) corresponding to the \( \sigma \)-fields \( \sigma(A_1) \) on \( \Omega \) and \( I_\gamma \) on \( J \). Let \( \mathcal{U} \subseteq \sigma(A_1) \otimes I_\gamma \) be a nonempty collection of tables, see 4.1, and let \( \alpha \) be a function from \( \mathcal{U} \times \Omega \) into the unit interval such that \( \alpha(U, \cdot) : \Omega \rightarrow [0,1] \) is Borel measurable for all \( U \in \mathcal{U} \). The pair \( (\mathcal{U}, \alpha) \) is called an inference rule with inferential function \( \alpha \) if \( \alpha \) is monotone, i.e.

\[
U_1 \subset U_2 \Rightarrow \alpha(U_1, \omega) \leq \alpha(U_2, \omega)
\]

for all \( U_1, U_2 \in \mathcal{U} \) and \( \omega \in \Omega \). In the sampling experiment the inferential function \( \alpha \) of the inference rule \( (\mathcal{U}, \alpha) \) is taken to be equal to the confidence level. Since in this case the trace depth is zero \( \alpha(U, \omega) \) is equal to the largest lower bound of the inference’s repeated sampling truth probability, i.e.

\[
\alpha(U, \omega) = \inf_{p \in J} p(U_p)
\]

see (4.2.1.2).

According to Berkum et al. (1996), in general there exists a collection \( \{ P^{A_1(\eta), \gamma} | \eta \in \Omega \} \) of inferential distributions corresponding to the inference rule \( (\mathcal{U}, \alpha) \), i.e. for all \( \eta \in \Omega \) \( P^{A_1(\eta), \gamma} \) is a probability measure on \( \mathcal{R}^{n, \gamma} := \sigma(\{U^\eta | U \in \mathcal{U}\}) \subseteq I_\gamma \) such that for all \( U \in \mathcal{U} \)
and there exists $V \in \mathcal{U}$ such that

$$U^\eta = V^\eta$$

and $P^A_1(\eta)^\gamma(U^\eta) = a(U, \eta)$.

see Definition 3.1 in Berkum et al. (1996).

This family of inferential distributions is uniquely determined by the inference rule $(U, a)$.

In this case the inferential distributions are called fiducial distributions as the inferential function is equal to the confidence level.

We now introduce the set $Q_{\omega, \gamma}$ of probability measures on $\Sigma|\mathcal{P}$ that are compatible with the ancillary information of $A_1$, i.e.

$$(4.2.1.1.2.1) \quad Q_{\omega, \gamma} := \{ q \in \Sigma|\mathcal{P} | q(\lambda(B)) = P^A_1(\omega)^\gamma(B) \text{ for all } B \in \mathcal{R}^\omega \gamma \},$$

where $\Sigma|\mathcal{P}$ is the set of probability measures on $\Sigma|\mathcal{P}$. For $q \in Q_{\omega, \gamma}$ the truth probability of the compound experiment as described in 4.2.1 is now equal to

$$\int \beta(p) q(dp),$$

and

$$(4.2.1.1.2.2) \quad \alpha_0(\gamma, U, \omega) = \inf_{q \in Q_{\omega, \gamma}} \int \beta(p) q(dp).$$

The confidence level $\alpha(U, \omega)$ of the table $U$ corresponding to the observation $\omega \in \Omega$ is defined to be

$$\alpha(U, \omega) = \inf_{\gamma \in \Gamma} \alpha_0(\gamma, U, \omega),$$

according to (4.2.3).

We now have determined the confidence level of a table that satisfies the condition that for every incidental experiment the trace depth equals zero. If the determination of the confidence level of a table leads to incidental experiments with tables satisfying the above condition, then of course the confidence level can be calculated also. The inferential function of inference rules in incidental experiments is taken to be equal to the confidence level. The corresponding family of inferential distributions is called a family of fiducial distributions. For more complex ancillary structures the confidence level of a table is determined by recursive application of this process.
4.2.2 Mainsteps in the calculation of the confidence level

Consider an experiment with trace depth \( k(\omega) = n, \ n \geq 0 \), for the observation \( \omega \). The confidence level \( \alpha(U, \omega) \) corresponding to a table \( U \) and the observation \( \omega \) is determined in the following way.

First we determine the truth probabilities \( \alpha_n(\gamma, U, \omega) \) in the \( n^{th} \) subexperiment for all \( \gamma \in \Gamma \). The trace depth in the \( n^{th} \) subexperiment is zero. So, according to (4.2.1.1)

\[
(4.2.2.1) \quad \alpha_n(\gamma, U, \omega) = \inf_{p \in \Gamma} \psi_n(p)(U_p) .
\]

Now suppose that we have calculated the truth probabilities \( \alpha_j(\gamma, U, \omega) \) in the \( j^{th} \) subexperiment for all \( \gamma \in \Gamma \) and all \( i + 1 \leq j \leq n \), where \( 0 \leq i \leq n - 1 \). The truth probability \( \alpha_i(\gamma, U, \omega) \) in the \( i^{th} \) subexperiment is, according to the theory in this section, determined as follows.

First determine for \( p \in R_\gamma \) the function

\[
(4.2.2.2) \quad \beta_i(p) = \psi_{i+1}(p)(U_p \cap \text{Ref}(\omega)) + \sum_{j=i+1}^{n-1} \psi_{i+1}(p)(C_j(\omega) \setminus C_{j+1}(\omega)) \alpha_j(\gamma, U, \omega) ;
\]

compare (4.2.1.3). Subsequently determine according to 4.2.1.1 the incidental experiment corresponding to the ancillary statistic \( A_{i+1} \) on \( (\Omega, \Sigma_i) \). We now have

\[
(4.2.2.3) \quad \alpha_i(\gamma, U, \omega) = \inf_{q \in \Gamma_{\omega, \gamma}} \int_{R_\gamma} \beta_i(p)q(dp) ;
\]

compare (4.2.1.1.2.1-2).

Remark. Suppose we have

\[
(4.2.2.4) \quad \psi_i(p)(C_i(\omega) \setminus C_{i+1}(\omega)) = 1
\]

for all \( p \in \mathcal{P} \) and \( 0 \leq i \leq n - 1 \). In this case we have for all \( p \in R_\gamma \)

\[
\beta_i(p) = \begin{cases} 
\alpha_{i+1}(\gamma, U, \omega) , & 0 \leq i < n - 1 , \\
\psi_n(p)(U_p) , & i = n - 1 .
\end{cases}
\]

Hence, for \( 1 \leq i \leq n - 1 \) we have

\[
(4.2.2.5) \quad \alpha_i(\gamma, U, \omega) = \alpha_0(\gamma, U, \omega) .
\]

and

\[
(4.2.2.6) \quad \alpha(U, \omega) = \inf_{\gamma \in \Gamma} \alpha_{n-1}(\gamma, U, \omega) .
\]
5 Examples

In this section we give some examples. In some of these examples both the set $\Omega$ of the sample space $(\Omega, \Sigma)$ and the set $\mathcal{P}$ of the probability model $(\mathcal{P}, \Sigma | \mathcal{P})$ are finite. We extensively discuss these elementary examples because in this way the basic ideas and concepts of our approach are illustrated preeminently. First we present the notation that we shall use. The inference model is represented by a matrix as shown below.

$$
\begin{array}{c|cccc}
\mathcal{P} & K_1 & & & \\
\downarrow & K_2 & & & \\
& \vdots & & & \\
L_1 & & & & \\
L_2 & & & & \\
\vdots & & & & \\
\Omega & 1 & 2 & 3 & \cdots \\
\end{array}
$$

In this picture the elements of the sample space $\Omega$ are denoted by $1, 2, 3, \ldots$, and the probability measures in $\mathcal{P}$ are indicated by $K_1, K_2, \ldots, L_1, L_2, \ldots$. The partition of $\mathcal{P}$ generating the $\sigma$-field $\mathcal{R}$ of interest is specified by bars in the picture, i.e.

$$
\mathcal{R} = \sigma(\{K_1, K_2, \ldots\}, \{L_1, L_2, \ldots\}.
$$

By use of the matrix $M$ we can compute probabilities as follows. The probability of the event $\{j\}$ corresponding to a probability measure $p \in \mathcal{P}$ is equal to

$$
p(\{j\}) = m_{ij} / \sum_{j \in \Omega} m_{ij}.
$$

where $i$ is the number of the row that corresponds to the probability measure $p$.

Example 5.1

The inference model

$$(\Omega, \Sigma = P(\Omega)), (\mathcal{P}, \Sigma | \mathcal{P}), \mathcal{R} \subset \Sigma | \mathcal{P}$$

is given by $(P(\Omega)$ is the powerset of $\Omega)$
The partition as described in (4.1) is written as

\[ \mathcal{P} = R_K \cup R_L , \]

where

\[ R_K := \{ K_1, K_2, K_3 \} \in \mathcal{R} , \]

\[ R_L := \{ L_1, L_2, L_3 \} \in \mathcal{R} . \]

So \( \Gamma := \{ K, L \} \).

The statistic

\[ I(i) = \begin{cases} 1, & i = 1 , \\ 2, & i = 2,3 \end{cases} \]

is the unique invariant statistic on \((\Omega, \Sigma)\). After the invariant reduction we obtain a reference model. So for every \( \omega \in \Omega \) the trace depth \( k(\omega) = 0 \) and \( \text{Ref}(\omega) = \Omega \). Furthermore

\[ \Sigma_0 = \sigma(\{1\}, \{2,3\}) \subset \Sigma \]

and for \( p \in \mathcal{P} \) we have that \( \psi_0(p) \) is the marginal distribution on \( \Sigma_0 \) corresponding to \( p \in \mathcal{P} \).

The reference model

\[ (\Omega, \Sigma_0 ) , (\mathcal{P}_0, \Sigma_0|\mathcal{P}_0) , \mathcal{R}_0 \subset \Sigma_0|\mathcal{P}_0 , \]

is depicted below

<table>
<thead>
<tr>
<th>( \psi_0(K_1) )</th>
<th>5</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi_0(K_2) )</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>( \psi_0(L_1) )</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>( \psi_0(L_2) )</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>( \psi_0(L_3) )</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>

\[ (5.1.1) \]

\[ \psi_0(K_2) = \psi_0(K_3) \]

We now specify a table \( U \in \Sigma \otimes \mathcal{R} \) by

\[ \begin{array}{c|ccc} K_1 & 10 & 15 & 15 \\ K_2 & 16 & 20 & 4 \\ K_3 & 16 & 4 & 20 \\ \hline L_1 & 20 & 10 & 10 \\ L_2 & 30 & 5 & 5 \\ L_3 & 24 & 8 & 8 \\ \hline 1 & 2 & 3 \end{array} \]
see (4.1.1) and (4.1.3). Use (4.2.1.2) to obtain the confidence level \( \alpha(U, \omega) \) of the table \( U \) and corresponding to the observation \( \omega \in \Omega \). We get

\[
\alpha(U, \omega) = \inf \{ \alpha_0(K, U, \omega), \alpha_0(L, U, \omega) \}.
\]

Use (4.2.1.1) to write

\[
\begin{align*}
\alpha_0(K, U, \omega) &= \inf \left\{ \frac{3}{4}, \frac{3}{5} \right\} = \frac{3}{5}, \\
\alpha_0(L, U, \omega) &= \inf \left\{ \frac{1}{2}, \frac{3}{4}, \frac{3}{5} \right\} = \frac{1}{2}.
\end{align*}
\]

(5.1.2)

So for all \( \omega \in \Omega \) we get \( \alpha(U, \omega) = \frac{1}{2} \).

**Example 5.2**

The inference model

\[ (\Omega, \Sigma = P(\Omega)), \ (\mathcal{P}, \overline{\Sigma}|\mathcal{P}). \quad \mathcal{R} \subset \overline{\Sigma}|\mathcal{P} \]

is given by

\[
\begin{array}{c|cccc}
\hline
 & K_1 & K_2 & L_1 & L_2 & L_3 \\
\hline
K_1 & 2 & 6 & 8 & 4 \\
K_2 & 4 & 6 & 4 & 6 \\
\hline
L_1 & 4 & 4 & 6 & 6 \\
L_2 & 6 & 2 & 4 & 8 \\
L_3 & 6 & 4 & 6 & 4 \\
\hline
 & 1 & 2 & 3 & 4
\end{array}
\]

(5.2.1)

The partition as described in (4.1) is written as

\[ \mathcal{P} = R_K \cup R_L, \]

where

\[ R_K := \{ K_1, K_2 \} \in \mathcal{R}, \]

\[ R_L := \{ L_1, L_2, L_3 \} \in \mathcal{R}. \]

So \( \Gamma := \{ K, L \}. \)

We now specify the table \( U \in \Sigma \cap \mathcal{R} \) by

\[
(5.2.2) \quad U^1 := R_L, \ U^2 := U, \ U^3 := R_K.
\]

see (4.1.1). All invariant and partially sufficient statistics on \((\Omega, \Sigma)\) are trivial. Hence,
\[ \Sigma_0 = \Sigma, \quad \mathcal{P}_0 = \mathcal{P}. \]

Fix \( \omega \in \Omega \), and take \( \omega := 1 \). Consider the statistic

\[
A_1(i) = \begin{cases} 
1, & i = 1, 2, \\
2, & i = 3, 4
\end{cases}
\]
on \((\Omega, \Sigma_0) = (\Omega, \Sigma)\). Let \( \psi(p) \) be the conditional probability distribution on \( \Sigma_0 = \Sigma \) given \( A_1 = A_1(\omega) \) corresponding to \( p \in \mathcal{P}_0 = \mathcal{P} \), and let \( \varphi(p) \) be the marginal probability distribution on

\[
\sigma(A_1) = \sigma(\{1, 2\}, \{3, 4\}) \subset \Sigma_0 = \Sigma
\]
corresponding to \( p \in \mathcal{P}_0 = \mathcal{P} \). We have

\[
\sigma(\psi) = \frac{\Sigma_0}{\mathcal{P}_0} = \frac{\Sigma}{\mathcal{P}},
\]

\[
\sigma(\varphi) = \sigma(\{K_1, L_1, L_2\}, \{K_2, L_3\}) \subset \frac{\Sigma_0}{\mathcal{P}_0} = \frac{\Sigma}{\mathcal{P}}.
\]

From (2.4,5) we conclude that \( A_1 \) is a partially ancillary statistic on \( (\Omega, \Sigma_0) \) with respect to \( \omega \in \Omega \). The statistic \( A_1 \) is also maximal and unique; see (2.6). After conditioning on \( A_1 = A_1(\omega) \) we obtain a reference model, i.e. all invariant, partially sufficient and partially ancillary statistics on the sample space \( (\Omega, \Sigma_1) = (\Omega, \Sigma) \) are trivial. So the trace depth \( k(\omega) = k(1) = 1 \). Since \( \mathcal{P}_0 = \mathcal{P} \) and \( \Sigma_1 = \Sigma \) for the map

\[
\psi_1 : (\mathcal{P}, \frac{\Sigma}{\mathcal{P}}) \to (\frac{\Sigma_1}{\mathcal{P}_1}, \frac{\Sigma_0}{\mathcal{P}_0})
\]
in (3.5) we have \( \psi_1 = \psi \). The reference model

\[
(\Omega, \Sigma_1), (\mathcal{P}_1, \frac{\Sigma_1}{\mathcal{P}_1}), \mathcal{R}_1 \subset \frac{\Sigma_1}{\mathcal{P}_1}
\]
is depicted below

\[
\begin{array}{c|cccc}
\psi_1(K_1) & 5 & 15 & 0 & 0 \\
\psi_1(K_2) & 8 & 12 & 0 & 0 \\
\hline
\psi_1(L_1) & 10 & 10 & 0 & 0 \\
\psi_1(L_2) & 15 & 5 & 0 & 0 \\
\psi_1(L_3) & 12 & 8 & 0 & 0 \\
\hline
& 1 & 2 & 3 & 4
\end{array}
\]

(5.2.3)

see (4.2). According to (4.3) we write

\[
\{1, 2\} = \text{Ref}(\omega) = C_1(\omega) \subset C_0(\omega) = \Omega.
\]
Obviously (4.1.3) is satisfied for all \( p \in \mathcal{P} \) and for \( \omega \). It is easily verified that (4.1.3) holds for all \( (\eta, p) \in \Omega \times \mathcal{P} \). According to 4.2.2 we first have to determine the truth probability \( \alpha_1(\gamma, U, \omega) \) in the 1st subexperiment. The reference model (5.2.3) of the 1st subexperiment is isomorphic with the reference model (5.1.1) in Example 5.1. Furthermore, the table \( U(1) \), see (4.2.2), corresponds to the table in Example 5.1. So \( \alpha_1(\gamma, U, \omega) \) equals the truth probability (5.1.2) in the 0th subexperiment of Example 5.1. Hence

\[
\begin{align*}
\alpha_1(K, U, \omega) &= \frac{3}{5}, \\
\alpha_1(L, U, \omega) &= \frac{1}{2}.
\end{align*}
\]

We now calculate the truth probability \( \alpha_0(\gamma, U, \omega) \) in the 0th subexperiment in the case \( \gamma = L \). So following 4.2.1 we introduce the function \( \beta \) on \( R_\gamma \) as described in (4.2.1.3) and (4.2.2.2). In (4.2.2.2) we have \( i = 0, n = 1 \) and therefore

\[
\begin{align*}
\beta(L_1) &= \psi_1(L_1)(U_{L_1} \cap \text{Re}(\omega)) = \psi_1(L_1)(\{1, 4\} \cap \{1, 2\}) = \psi_1(L_1)(\{1\}) = \frac{1}{2} , \\
\beta(L_2) &= \psi_1(L_2)(\{1\}) = \frac{3}{4} , \\
\beta(L_3) &= \psi_1(L_3)(\{1\}) = \frac{2}{5} .
\end{align*}
\]

We now describe the incidental experiment corresponding to \( A_1 \) and \( \gamma = L \in \Gamma \); see 4.2.1.1. The sample space \( (\Omega, \sigma(A_1)) \) and the probability model \( (J, \sigma(A_1)|J) \), see (4.2.1.1.4), of the incidental experiment are given below

\[
\begin{array}{c|cc}
\varphi(L_1) & 4 & 6 \\
\varphi(L_3) & 5 & 5 \\
\hline
1, 2 & 3, 4
\end{array}
\]

Note that in this case (see (4.2.1.1.2))

\[
\lambda(L_i) = \varphi \psi_0(L_i) = \varphi(L_i), \quad i = 1, 2, 3.
\]

We now specify the interest in the incidental experiment. It follows from (5.2.5) and (5.2.6)

\[
\begin{align*}
\sigma(\lambda)|R_\gamma &= \sigma(\{L_1, L_2\}, \{L_3\}) \subset \Sigma|R_\gamma , \\
\sigma(\beta) &= \sigma(\{L_1\}, \{L_2\}, \{L_3\}) = \Sigma|R_\gamma .
\end{align*}
\]

According to the definition in 4.2.1.1.1 \( \sigma(\lambda)|R_\gamma \) is the unique sub-\( \sigma \)-field of \( \sigma(\lambda)|R_\gamma \) which is crosscoherent with \( \sigma(\beta) \). So the \( \sigma \)-field \( I_\gamma \subset \sigma(A_1)|J \) of interest in the incidental experiment is given by

\[
I_\gamma = \sigma(A_1)|J .
\]

The inference model

23
for the incidental experiment is also a reference model. So the SCIRA scheme is trivial in this case. Therefore the confidence level of tables for the incidental experiment can be calculated according to (4.2.1.2). We now describe a family of fiducial distributions, see 4.2.1.1.2, as a family of appropriate weighting distributions for \( \beta \) on \( R_y \); see (5.2.5). To construct such a family of fiducial distributions we first introduce an inference rule containing three tables \( U_1, U_2, U_3 \in \sigma(A_1) \otimes I_7 \). These tables are specified by (see (4.1.1))

\[
(U_1)^{1,2} = (U_2)^{1,2} := \{ \varphi(L_3) \}, \quad (U_3)^{1,2} := J,
\]

(5.2.7)

\[
(U_1)^{3,4} := \{ \varphi(L_1) \}, \quad (U_2)^{3,4} = (U_3)^{3,4} := J.
\]

Since the value of the inferential function \( \alpha \) of the inference rule has to be taken equal to the confidence level, we conclude that for \( i = 1, 2, 3 \) and all \( \eta \in \Omega \) we have

\[
\alpha(U_i, \eta) = \inf_{p \in J} \{ p((U_i)_p) \}.
\]

see (4.2.1.2). For all \( \eta \in \Omega \) we obtain

\[
\alpha(U_1, \eta) = \frac{1}{2}, \quad \alpha(U_2, \eta) = \frac{2}{3}, \quad \alpha(U_3, \eta) = 1.
\]

According to Berkum et al. (1996) there exists a unique family of fiducial distributions on \( I_7 \); see 4.2.1.1.2. The fiducial distribution

\[
P^\omega_{A_1}(\omega, \gamma)
\]

on \( I_7 \) is calculated by use of Corollary 1.1 in Berkum et al. (1996). We obtain

\[
P^\omega_{A_1}(\omega, \gamma)(\{ \varphi(L_1) \}) = \frac{2}{3}, \quad P^\omega_{A_1}(\omega, \gamma)(\{ \varphi(L_3) \}) = \frac{2}{3}.
\]

The set \( Q_{\omega, \gamma} \) in (4.2.1.1.2.1) can be put in the form

\[
Q_{\omega, \gamma} := \{ p_x \in \widehat{\Sigma|\mathcal{P}} \mid 0 \leq x \leq \frac{2}{3} \},
\]

where the probability distribution \( p_x \) on \( \widehat{\Sigma|\mathcal{P}} \) is given by

\[
p_x(\{K_1\}) = p_x(\{K_2\}) = 0, \\
p_x(\{L_1\}) = x, \quad p_x(\{L_2\}) = \frac{2}{3} - x, \quad p_x(\{L_3\}) = \frac{2}{3}.
\]

From (4.2.1.1.2.2) and (5.2.5) for the table \( U \) in (5.2.2) we get
\[ \alpha_0(\gamma, U, \omega) = \alpha_0(L, U, 1) = \inf \left\{ \frac{1}{2} x + \frac{3}{4} \left( \frac{2}{5} - x \right) + \frac{3}{5} \cdot \frac{3}{5} \mid 0 \leq x \leq \frac{2}{5} \right\} = \frac{14}{25}, \]

and by similar calculations we obtain

\[ \alpha_0(K, U, 1) = \frac{32}{60}. \]

The confidence level \( \alpha(U, \omega) \) of the table \( U \) and corresponding to the observation \( \omega = 1 \) is evaluated following definition (4.2.1.1.2.3). We have

\[ \alpha(U, \omega) = \alpha(U, 1) = \frac{14}{25}. \]

Obviously \( \alpha(U, 2) = \alpha(U, 1) = \frac{14}{25} \), and by use of the same procedure one obtains

\[ \alpha(U, 3) = \alpha(U, 4) = \frac{9}{20}. \]

If we calculate the confidence level \( \alpha \) of this table \( U \) using the argument of repeated sampling only, then obviously we get \( \alpha = \frac{1}{2} \). As a consequence of the verification process of statistical methods as described in this paper, the confidence level \( \alpha \) of the table \( U \) depends on the observation. We got

\[ \alpha(U, 1) = \alpha(U, 2) = \frac{14}{25}, \]

\[ \alpha(U, 3) = \alpha(U, 4) = \frac{9}{20}. \]

**Example 5.3**

The inference model

\[ (\Omega, \Sigma = P(\Omega)) , (\mathcal{P}, \overline{\Sigma}|\mathcal{P}), \mathcal{R} \subseteq \overline{\Sigma}|\mathcal{P} \]

is given by

\[
\begin{array}{c|cccccc}
K_1 & 1 & 3 & 4 & 2 & 0 & 20 \\
K_2 & 4 & 6 & 4 & 6 & 4 & 6 \\
L_1 & 2 & 2 & 3 & 3 & 10 & 10 \\
L_2 & 6 & 2 & 4 & 8 & 6 & 4 \\
L_3 & 3 & 2 & 3 & 2 & 20 & 0 \\
L_4 & 3 & 1 & 2 & 4 & 12 & 8 \\
\hline
1 & 2 & 3 & 4 & 5 & 6
\end{array}
\]

(5.3.1)

The partition as described in (4.1) is written as

\[ \mathcal{P} = R_K \cup R_L \]

where

25
\[ R_K := \{K_1, K_2\} \subset \mathcal{R} \]
\[ R_L := \{L_1, L_2, L_3, L_4\} \subset \mathcal{R} \]

So \( \Gamma = \{K, L\} \).

We now specify a table \( U \in \Sigma \otimes \mathcal{R} \) by

\[
U^1 = U^4 = U^5 := R_L ,
\]
\[
U^2 = U^3 = U^6 := R_K ;
\]

see (4.1.1). All invariant and partially sufficient statistics on \((\Omega, \Sigma)\) are trivial. Hence,

\[
\Sigma_0 = \Sigma , \quad \mathcal{P}_0 = \mathcal{P} .
\]

Fix \( \omega \in \Omega \), and take \( \omega := 1 \). Consider the statistic

\[
A_1(i) = \begin{cases} 1, & i = 1, 2, 3, 4 \, , \\ 2, & i = 5, 6 \end{cases}
\]

on \((\Omega, \Sigma_0) = (\Omega, \Sigma)\). Let \( \psi(p) \) be the conditional probability distribution on \( \Sigma_0 = \Sigma \) given \( A_1 = A_1(\omega) \) corresponding to \( p \in \mathcal{P}_0 = \mathcal{P} \), and let \( \varphi(p) \) be the marginal probability distribution on

\[
\sigma(A_1) = \sigma(\{1, 2, 3, 4\}, \{5, 6\}) \subset \Sigma_0 = \Sigma
\]
corresponding to \( p \in \mathcal{P}_0 = \mathcal{P} \). We have

\[
\sigma(\psi) = \sigma(\{K_1\}, \{K_2\}, \{L_1\}, \{L_2, L_4\}, \{L_3\}) \subset \Sigma_0 | \mathcal{P}_0 = \Sigma | \mathcal{P} ,
\]
\[
\sigma(\varphi) = \sigma(\{K_1, L_1, L_3, L_4\}, \{K_2, L_2\}) \subset \Sigma_0 | \mathcal{P}_0 = \Sigma | \mathcal{P} .
\]

From (2.4,5) we conclude that \( A_1 \) is a partially ancillary statistic on \((\Omega, \Sigma_0)\) with respect to \( \omega \in \Omega \). The statistic \( A_1 \) is also maximal and unique; see (2.6). After conditioning on \( A_1 = A_1(\omega) \) we obtain an inference model such that all invariant and partially sufficient statistics on its sample space \((\Omega, \Sigma_0) = (\Omega, \Sigma)\) are trivial. Hence,

\[
\Sigma_1 = \Sigma_0 = \Sigma ;
\]

see Section 3. Since \( \mathcal{P}_0 = \mathcal{P} \) and \( \Sigma_1 = \Sigma \) for the map

\[
\psi_1 : (\mathcal{P}, \Sigma | \mathcal{P}) \rightarrow (\hat{\Sigma}_1, \Sigma_1)
\]

in (3.5) we have \( \psi_1 = \psi \). The new inference model
The inference model (5.3.3) of the 1\textsuperscript{st} subexperiment is isomorphic with the inference model (5.2.1) in Example 5.2. So the statistic \( A_2 \) on \( (\Omega, \Sigma_1) \) defined by

\[
A_2(i) = \begin{cases} 
1, & i = 1, 2, \\
2, & i = 3, 4, \\
3, & i = 5, 6
\end{cases}
\]

is the unique and maximal partially ancillary statistic on \( (\Omega, \Sigma_1) \) with respect to \( \omega \in \Omega \); compare the statistic \( A_1 \) in Example 5.2. After conditioning on \( A_2 = A_2(\omega) \) we obtain a reference model. So the trace depth

\[
k(\omega) = k(1) = 2,
\]

\[
\Sigma_2 = \Sigma_1 = \Sigma_0 = \Sigma.
\]

This reference model is depicted below

\[
\begin{array}{c|cccccc}
\psi_2(K_1) & 5 & 15 & 0 & 0 & 0 & 0 \\
\psi_2(K_2) & 8 & 12 & 0 & 0 & 0 & 0 \\
\psi_2(L_1) & 10 & 10 & 0 & 0 & 0 & 0 \\
\psi_2(L_2) & 15 & 5 & 0 & 0 & 0 & 0 \\
\psi_2(L_3) & 12 & 8 & 0 & 0 & 0 & 0 \\
\end{array}
\]

\[
\psi_2(L_2) = \psi_2(L_4)
\]

\[
(5.3.4)
\]

Here

\[
\psi_2 : (\mathcal{P}, \Sigma_2 | \mathcal{P}) \rightarrow (\bar{\Sigma}_2, \bar{\Sigma}_2)
\]

is a map such that \( \psi_2(p), p \in \mathcal{P} \), is the conditional probability distribution on \( \Sigma_2 \) given \( A_2 = A_2(\omega) \) and corresponding to \( p \in \mathcal{P} \). For \( \omega = 1 \in \Omega \) we constructed the inference models

\[
(\Omega, \Sigma_i) : (\mathcal{P}, \Sigma_i | \mathcal{P}_i), \quad \mathcal{R}_i \subset \Sigma_i | \mathcal{P}_i, \quad 0 \leq i \leq k(\omega) = 2.
\]

corresponding to the partially ancillary trace
\[ \{1, 2\} = \text{Ref}(\omega) = C_2(\omega) \subseteq C_1(\omega) = \{1, 2, 3, 4\} \subseteq C_0(\omega) = \Omega; \]

see (3.9-11). Obviously (4.1.3) is satisfied for all \( p \in \mathcal{P} \) and for \( \omega \). It is easily verified that (4.1.3) holds for all \((\eta, p) \in \Omega \times \mathcal{P}\). It follows from the isomorphism between the inference models (5.3.3) and (5.2.1) and the correspondence of the tables involved, that the truth probabilities \( \alpha_2(\gamma, U, \omega) \) and \( \alpha_1(\gamma, U, \omega) \) are equal to the truth probabilities of the 1\(^{\text{st}}\) and 0\(^{\text{th}}\) subexperiment in Example 5.2. Hence,

\[
\alpha_2(L, U, \omega) = \frac{1}{2}, \quad \alpha_1(L, U, \omega) = \frac{14}{25}.
\]

We now calculate the truth probability \( \alpha_0(\gamma, U, \omega) \) in the 0\(^{\text{th}}\) subexperiment in the case \( \gamma = L \). Following 4.2.1 we introduce the function \( \beta \) on \( R_\gamma \) as described in (4.2.1.3) and (4.2.2.2). In (4.2.2.2) we have \( i = 0, n = 2 \) and therefore

\[
\beta(L_1) = \psi_1(L_1)(U_1 \cap \text{Ref}(\omega)) + \psi_1(L_1)(C_1(\omega) \setminus C_2(\omega)) \alpha_1(L, U, \omega) = \\
= \psi_1(L_1)(\{1\}) + \frac{14}{25} \psi_1(L_1)(\{3, 4\}) = \frac{23}{25}.
\]

(5.3.5)

\[
\beta(L_2) = \beta(L_4) = \frac{159}{250}.
\]

\[
\beta(L_3) = \frac{145}{250}.
\]

We now describe the incidental experiment corresponding to \( A_1 \) and \( \gamma = L \in \Gamma; \) see 4.2.1.1. The sample space \((\Omega, \sigma(A_1))\) and the probability model \((J, \sigma(A_1)|J)\), see (4.2.1.1.4), of the incidental experiment are given below

<table>
<thead>
<tr>
<th>( \varphi(L_1) )</th>
<th>( \varphi(L_2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

\[(\ref{5.3.6}) \quad \varphi(L_1) = \varphi(L_3) = \varphi(L_4) \]

Note that in this case (see (4.2.1.1.2))

\[
\lambda(L_i) = \varphi_\gamma(L_i) = \varphi(L_i), \quad i = 1, 2, 3, 4.
\]

We now specify the interest in the incidental experiment. It follows from (5.3.5) and (5.3.6)

\[
\sigma(\lambda)|R_\gamma = \sigma(\{L_1, L_3, L_4\}, \{L_2\}) \subseteq \Sigma|R_\gamma,
\]

\[
\sigma(\beta) = \sigma(\{L_1\}, \{L_2, L_4\}, \{L_3\}) \subseteq \Sigma|R_\gamma.
\]

According to the definition in 4.2.1.1.1 \( \sigma(\lambda)|R_\gamma \) is the unique sub-\( \sigma \)-field of \( \sigma(\lambda)|R_\gamma \) which is crosscoherent with \( \sigma(\beta) \). So the \( \sigma \)-field \( I_\gamma \subseteq \sigma(A_1)|J \) of interest in the incidental experiment is given by

\[
I_\gamma = \sigma(A_1)|J.
\]

28
The inference model

\[
(\Omega, \sigma(A_1)) , (J, \sigma(A_1)||J) , I_\gamma = \sigma(A_1)||J
\]

for the incidental experiment is also a reference model. So the SCIRA scheme is trivial in this case. Following a similar procedure as in Example 5.2 we obtain the following fiducial distribution on \( I_\gamma \)

\[
P^{A_1(\omega), \gamma} (\{\varphi(L_1)\}) = \frac{1}{3} ,
\]

\[
P^{A_1(\omega), \gamma} (\{\varphi(L_2)\}) = \frac{2}{3} ,
\]

The set \( Q_{\omega, \gamma} \) in (4.2.1.1.2.1) can be put in the form

\[
Q_{\omega, \gamma} := \{ p_{x,y} \in \Sigma | \mathcal{P} \mid 0 \leq x, 0 \leq y, x + y \leq \frac{1}{3} \}
\]

where the probability distribution \( p_{x,y} \) on \( \Sigma | \mathcal{P} \) is given by

\[
p_{x,y}(\{K_1\}) = p_{x,y}(\{K_2\}) = 0 .
\]

\[
p_{x,y}(\{L_1\}) = x , \quad p_{x,y}(\{L_2\}) = \frac{2}{3} , \quad p_{x,y}(\{L_3\}) = y ,
\]

\[
p_{x,y}(\{L_4\}) = \frac{1}{3} - x - y .
\]

From (4.2.1.1.2.2) and (5.3.5) for the table \( U \) in (5.3.2) we get

\[
\alpha_0(\gamma, U, \omega) = \alpha_0(L, U, 1) = \\
= \inf \{ \frac{134}{250} x + \frac{159}{250} \cdot \frac{2}{3} + \frac{145}{250} , \ y + \frac{159}{250} (\frac{1}{3} - x - y) \mid 0 \leq x, 0 \leq y, x + y \leq \frac{1}{3} \} = \frac{226}{375} ,
\]

and by similar calculations we obtain

\[
\alpha_0(K, U, 1) = \frac{163}{250} .
\]

The confidence level \( \alpha(U, \omega) \) of the table \( U \) and corresponding to the observation \( \omega = 1 \) is evaluated following definition (4.2.1.1.2.3). We have

\[
\alpha(U, \omega) = \alpha(U, 1) = \frac{226}{375} ,
\]

Obviously \( \alpha(U, 2) = \alpha(U, 1) = \frac{226}{375} , \) and by use of the same procedure one obtains

\[
\alpha(U, 3) = \alpha(U, 4) = \frac{317}{606} ,
\]

\[
\alpha(U, 5) = \alpha(U, 6) = \frac{7}{10} .
\]
If we calculate the confidence level $\alpha$ of this table $U$ using the argument of repeated sampling only, then obviously we get $\alpha = \frac{1}{2}$. As a consequence of the verification process of statistical methods as described in this paper, the confidence level $\alpha$ of the table $U$ depends on the observation. We got

$$\alpha(U, 1) = \alpha(U, 2) = \frac{228}{600},$$

$$\alpha(U, 3) = \alpha(U, 4) = \frac{317}{600},$$

$$\alpha(U, 5) = \alpha(U, 6) = \frac{7}{10}.$$ 

**Example 5.4**

Consider an experiment that consists of two successive steps. Only if step 1 succeeds, step 2 is performed. Only if step 2 succeeds a normally distributed variable $X$ is observed with expected value $\mu$ and variance $\sigma^2$. The parameter of interest is $\mu$. Let

$$P(\text{step 1 succeeds}) = (1 + \sigma/|\mu|)^{-1}$$

and

$$P(\text{step 2 succeeds}) = (1 + \sigma)^{-1}.$$ 

The experiment is repeated $n$ times. Let $N_i$ denote the number of successful steps $i$. Now the resulting sufficiently reduced observation can be represented as $(n_1, n_2, \bar{X}, s)$, where $\bar{X}$ is the sample mean of the $n_2 \geq 1$ observations of $X$ and $s^2$ is the sample variance ($n_2 \geq 2$). There are two levels of ancillarity. The statistic $N_1$ is the maximal partially ancillary statistic on the first level ($n_1 > 0$), whereas $(N_2, S)$ is the similar statistic on the second level.

The reference set is characterized by $N_1 = n_1, N_2 = n_2$ and $S = s$. The verification of inference tables is very similar to that described in 0.1. Now the fiducial distribution for $\sigma$ represents the inference for $\sigma$ based upon the value $(n_2, s)$ and conditional distribution of $(N_2, S)$, given $N_1 = n_1$. Consider as an example the case that $n = n_1 = 3$. Inference tables $\sigma < \sigma_\alpha$ with confidence levels $\alpha$ can be defined. If $n_2 = 0$ then $\sigma_\alpha = \infty$. If $n_2 = 1$ we have

$$\sigma_\alpha = (\alpha^{-\frac{1}{3}} - 1)^{-1}.$$ 

For $n_2 = 2$ and $n_2 = 3 \sigma_\alpha$ is the value of $\sigma$ that satisfies the equations:

$$P(N_2 < 2 \mid N_1 = 3) + P(N_2 = 2 \mid N_1 = 3) \cdot P(S > s \mid N_2 = 2, N_1 = 3) = \alpha,$$

$$P(N_2 < 3 \mid N_1 = 3) + P(N_2 = 3 \mid N_1 = 3) \cdot P(S > s \mid N_2 = 3, N_1 = 3) = \alpha,$$

respectively. The fiducial probability of the parameter statement $\sigma < \sigma_\alpha$ equals $\alpha$ (Berkum et al. (1996)). If $n_2 = 3$ the implied unique fiducial distribution is easily seen to be given by

$$F(\sigma) := 1 - (1 - e^{-s^2/\sigma^2})/(1 + \sigma)^3.$$
If the inference for \( \sigma \) is based only upon the value \( s \) and conditional distribution of \( S \), given \( N_1 = 3 \) and \( N_2 = 3 \) (compare 0.1) the resulting fiducial distribution is given by

\[
G(\sigma) := e^{-s^2/\sigma^2}.
\]

Obviously, as \( F > G \), this inference for \( \sigma \) results in too conservative inferential statements for the parameter of interest \( \mu \).

**Example 5.5**

Consider the well-known Behrens-Fisher problem two independent random samples from two normal distributions with parameters \( (\mu_x, \sigma_x^2) \) and \( (\mu_y, \sigma_y^2) \). The parameter of interest is \( \mu_x - \mu_y \). The statistic \( (s_x^2, s_y^2) \) is a maximal partially ancillary statistic, so \( s_x^2 \) and \( s_y^2 \) are known characteristics of the reference set. Parameter statements depend on the value of \( x - y \). So, the \( \sigma \)-field \( \sigma(\beta) \) defined in 4.2.1.1.1 is determined by the possible values of

\[
\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y},
\]

the variance of \( x - y \), and it is clear that the \( \sigma \)-field of interest should be isomorphic with \( \sigma(\beta) \). The conditions of 4.2.1.1.1 for crosscoherence are satisfied if \( F_0 \) is chosen to be equal to \( \sigma(\beta) \) and \( F_1 \) is for example the \( \sigma \)-field corresponding to the marginal distributions of \( s_x^2 \).

The weaker condition as mentioned in 4.2.1.1.1 is also satisfied if \( F_0 \) and \( F_1 \) are taken to be the \( \sigma \)-fields that are determined by \( s_x^2 \) and \( s_y^2 \) respectively.

**References**


Welch, B.L. (1947). The generalization of ‘Student’s’ problem when several different population variances are involved, *Biometrika* 34, 28–35.