Large Sample Inference about the Ratio of Means in Two Independent Populations

Mohammad Reza Mahmoudi[*]

Department of Statistics, College of Science, Fasa University
Fas, Iran
mahmoudi.m.r@fasau.ac.ir

Javad Behboodian

Department of Statistics, Shiraz University
Shiraz, Iran, 71454
behboodian@susc.ac.ir

Mohsen Maleki

Department of Statistics, Shiraz University
Shiraz, Iran, 71454
stud.stat@yahoo.com

Abstract

In a number of situations, for example in biology, psychology and neurosciences, researchers are interested in the ratio of two measured quantities. In this paper, we give an overview of different methods to constructing confidence limits for the ratios. Then by using the limiting theorems, a pivotal quantity for the ratio of population means will be presented. The results will be applied to construct confidence intervals and perform test of hypothesis. Finally, to investigate the ability of the proposed method, a modest simulation study is provided.

Keywords: Cramer's Theorem; Fieller's Theorem; Ratio of Means; Simulation; Slutsky's Theorem.

2000 Mathematics Subject Classification: 62F03; 62F05; 62F12; 62H20.

1. Introduction

In a number of situations, it is of interest to make inference about the ratio of means in two independent populations. For example, in biology, psychology and neurosciences, researchers are interested in the ratio of two measured quantities [See Refs. 1–4 for more details]. This parameter is more applicable than the difference of means in some applications. The advantage of using ratio instead of difference lies in the fact that the difference of two small means is also small and has no meaningful description. Fieller's theorem is the classic solution to constructing confidence limits for the ratios [See Refs. 5–7 for more details]. Most studies in psychology and the cognitive neurosciences require very specific assumptions about the distribution of

---

[*] Corresponding author.

numerator and denominator of the ratio [See Refs. 8–9 for more details]. If these assumptions are not satisfied, the confidence limits are not reliable. Other studies have applied appropriate methods such as Fieller, Taylor and special bootstrap methods which we will discussed later.

This paper organized as follows: An overview of alternatives to constructing confidence limits for the ratios will be studied in Section 2. In Section 3, by using the limiting theorems, a pivotal quantity for the ratio of population means will be presented. It will be applied to construct confidence intervals and to perform test of hypothesis. Finally, to investigate the ability of the proposed method, a simulation study is provided in Section 4.

2. Preliminaries and Different Approaches

Let X and Y be two independent random variables with means μ_1 and μ_2 and unknown variance σ^2. Assume $X_1, \dots, X_{n_1}$ and $Y_1, \dots, Y_{n_2}$ are two independent samples from X and Y, respectively. We are interested to make inference about the parameter $\mu = \frac{\mu_1}{\mu_2}$ ($\mu_2 \neq 0$). Tamhane and Logan [10] considered tests of the null-hypothesis $H_0: \mu = \mu_0$, where μ_0 is an hypothesized value of μ.

Fieller-type $(1-\alpha)$ interval [See Refs. 7, 11 for more details] is presented by

$$\frac{B}{-2A} \pm \frac{\sqrt{(B/2)^2 - AC}}{A}, \qquad (2.1)$$

where

$A = (S_P^2 C^2 / n_1) - \bar{X}^2$, $B = 2\bar{X}\bar{Y}$, $C = (S_P^2 c^2 / n_2) - \bar{Y}^2$, $c = t_{1-\alpha/2}(df)$, S_P^2 is sample pooled variance and $df = \frac{S_P^2/n_2 + \mu^2 S_P^2/n_1}{S_P^2/(n_2{}^2(n_2-1)) + \mu^4 S_P^2/(n_1{}^2(n_1-1))}$.

As can be seen, the confidence limits obtained from the Fieller-type interval depends on the degrees of freedom (df) which depends on the unknown parameter μ. In the test of hypothesis, the degrees of freedom is computed under the fixed value $\mu = \mu_0$. For confidence interval construction we can plug-in the estimated ratio parameter $\hat{\mu} = \frac{\bar{X}}{\bar{Y}}$ to compute the degrees of freedom [See Ref. 12 for more details]. If the denominator $(\bar{Y})$ is significantly different from zero at a significance level of α, then we get bounded confidence intervals (CIs). But if the denominator is close to zero, the confidence interval (CI) is unbounded.

2.1. Alternative approaches

In this subsection we give an overview of alternatives to Fieller's method which are applied in the statistical literature and other studies.

2.1.1. Taylor method

The Taylor method measures a linear approximation for the sample estimates. The approximate confidence limits for the Taylor method are

$$\hat{\mu} \pm t_{1-\alpha/2}(df)|\hat{\mu}| \sqrt{S_P^2 \left(\frac{1}{n_1 \bar{X}^2} + \frac{1}{n_2 \bar{Y}^2} \right)}. \qquad (2.2)$$

Although the Taylor approximation has good result, but, like Fieller's method the approximation fails, when the denominator is not significantly different from zero. But, if the denominator has small coefficients of variation (CV), the Taylor method provides a serious alternative to the Fieller's method [See Refs. 11, 13–14 for more details].

2.1.2. Bootstrap methods

The bootstrap method is a general method which allows constructing confidence limits in an easy and consistent way, even for very complicated statistic [See Ref. 15 for more details]. It uses the measured samples as a basis for resampling with the aim of creating an approximation to the population distribution. In order to determine confidence limits, we draw a large number of samples (with replacement and every sample has the same size as the original sample) with bootstrap methods. Then we calculate the ratio $\frac{\bar{X}}{\bar{Y}}$ for each sample. The empirical distribution of these resampled ratios is used for determining the CIs. In the standard bootstrap method, the CIs are the $\alpha/2$ and $(1-\alpha/2)$ percentiles of the empirical distribution. Bootstrap methods can provide an alternative to methods, especially in cases where (X,Y) is not normally distributed [See Refs. 8–9, 16, 11 for more details]. However, standard bootstrap methods face two problems when dealing with ratios [See Refs. 17–18 for more details]: (a) Bootstrap confidence limits can be erroneous if the variance of the statistic does not exist as in the case of ratios. (b) Bootstrap confidence limits can be arbitrary large deviations from the intended confidence level for ratios. Hwang (1995) solved these problems by applying a special bootstrap method [See Ref. 19 for more details]. He performed the bootstrap on the

$$T_0 = \frac{\bar{X} - \mu \bar{Y}}{\sqrt{S_P^2 \left(\frac{1}{n_2} + \mu^2 \frac{1}{n_1} \right)}}. \tag{2.3}$$

The method first uses the bootstrap samples to determine the $(1-\alpha)$ quantiles of T_0 and then proceeds as the Fieller's method does. Therefore, the Hwang bootstrap is more general and better than the standard bootstrap method.

3. Large Samples Inference

In this section, we give our asymptotic results. We will use a methodology similar to that was applied in [20-24]. Following theorem that is the main theorem of this article, will be needed to make inference about the parameter μ.

Theorem 1. *Under the assumptions,*

$$\sqrt{n} \left(\frac{\bar{X} - \mu \bar{Y}}{\sqrt{\sigma^2 (1 + \mu^2)}} \right) \xrightarrow{\mathcal{L}} N(0,2), \qquad as \ n \to \infty \ ,$$

where $n = min(n_1, n_2)$.

Proof. By using the Central Limit Theorem, we have,

$$\sqrt{n_1} \left(\frac{\bar{X} - \mu_1}{\sigma} \right) \xrightarrow{\mathcal{L}} N(0,1),$$

and

$$\sqrt{n_2}\left(\frac{\bar{Y}-\mu_2}{\sigma}\right) \xrightarrow{\mathcal{L}} N(0,1),$$

as $n \to \infty$.

Since the samples are independent, Slutsky's theorem [See Ref. 25 for more details] gives

$$\sqrt{n}\begin{pmatrix} \dfrac{\bar{X}-\mu_1}{\sigma} \\ \dfrac{\bar{Y}-\mu_2}{\sigma} \end{pmatrix} \xrightarrow{\mathcal{L}} N\left(\mathbf{0}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right), \, as \, n \to \infty.$$

Now define $f: \mathbb{R}^2 \to \mathbb{R}$ as $f(x_1, x_2) = x_1 - \mu x_2$. Then the gradient function with respect to f is $\nabla f(x_1, x_2) = (1, -\mu)$. Since ∇f is continuous, therefore by Cramer's rule we have

$$\sqrt{n}\left(f\left(\frac{\bar{X}-\mu_1}{\sigma}, \frac{\bar{Y}-\mu_2}{\sigma}\right)\right) = \sqrt{n}\left(\frac{\bar{X}-\mu\bar{Y}}{\sqrt{\sigma^2(1+\mu^2)}}\right) \xrightarrow{\mathcal{L}} N(0,2), \quad as \, n \to \infty. \quad \blacksquare$$

So we have just proved

$$T_n = \sqrt{n}\left(\frac{\bar{X}-\mu\bar{Y}}{\sqrt{2\sigma^2(1+\mu^2)}}\right) \xrightarrow{\mathcal{L}} N(0,1), \quad as \, n \to \infty. \tag{3.1}$$

In Sections 3.1 and 3.2, this result is used to construct an asymptotic CI and carry out a test of hypothesis.

3.1. Asymptotic confidence interval

Note that T_n given by Eq. (3.1) depends on the unknown parameters μ and σ^2. So it cannot be used as a pivotal quantity for the parameter μ.

Theorem 2. *Under the same assumptions as Theorem 1,*

$$T_n^* = \sqrt{n}\left(\frac{\bar{X}-\mu\bar{Y}}{\sqrt{2S_P^2(1+\hat{\mu}^2)}}\right) \xrightarrow{\mathcal{L}} N(0,1), \quad as \, n \to \infty. \tag{3.2}$$

Proof. By the Weak Law of Large Numbers, we have

$$S_P^2 \xrightarrow{p} \sigma^2, and \, \hat{\mu} \xrightarrow{p} \mu \, as \, n \to \infty.$$

The proof is now completed by using Theorem 1 and Slutsky's Theorem. $\quad \blacksquare$

Now, T_n^* can be used as a pivotal quantity to construct asymptotic CI for μ,

$$\frac{\bar{X} \pm Z_{1-\alpha/2}\sqrt{2S_P^2(1+\hat{\mu}^2)/n}}{\bar{Y}}. \tag{3.3}$$

3.2. Hypothesis testing

Hypothesis testing about μ is important in practice. For instance, the assumption $\mu = 1$ is equivalent to the assumption $\mu_1 = \mu_2$. In general, to test $H_0: \mu = \mu_0$, the test statistic can be

$$T_0^* = \sqrt{n}\left(\frac{\bar{X} - \mu_0\bar{Y}}{\sqrt{2S_P^2(1 + \mu_0^2)}}\right) . \qquad (3.4)$$

By similar methodology which was applied in Theorem 2, it can be shown that under null hypothesis, T_0^* has asymptotic standard normal distribution.

4. Simulation Study

In this section, numerous data sets are generated and analyzed to investigate the performance of proposed method, for different values of (n_1, n_2) and μ. We investigate that Eq. (3.3) and Eq. (3.4) are actually the asymptotic CI and hypothesis test statistic. For each parameter setting, the empirical coverage probability (percent of runs which Eq. (3.3) contains true μ) is estimated based on 1000 simulation runs. Also for each run, the value of test statistic in Eq. (3.4) is computed and then normal Q–Q plots of the test statistic (3.4) are plotted. The Shapiro-Wilk's normality test also is applied to verify normality of test statistic (3.4)

Example 4.1. Exponential populations

We assume X and Y have Exponential distribution with mean μ_1 and μ_2, respectively. The empirical coverage probabilities for different parameter setting are summarized in Table 1.

The results reveal that the empirical coverage probability of proposed method is very close to the nominal level (0.95) as sample size is growing, and so the type I error are asymptotically controlled by method. In other words, we can accept Eq. (3.3) is the asymptotic CI for μ. Also Figure 1 and Table 2 show the Q–Q plots against the standard normal distribution and the results of Shapiro-Wilk's normality test for the test statistic in Eq. (3.4), respectively.

It confirmed that the asymptotic approximation seems to be quite satisfactory in all of the cases considered (P-Value is more than 0.05). Therefore our approach is a good alternative to construct a CI and perform a test of hypothesis for the ratio of means in two independent populations.

Table 1. The empirical coverage probability (Exponential populations)

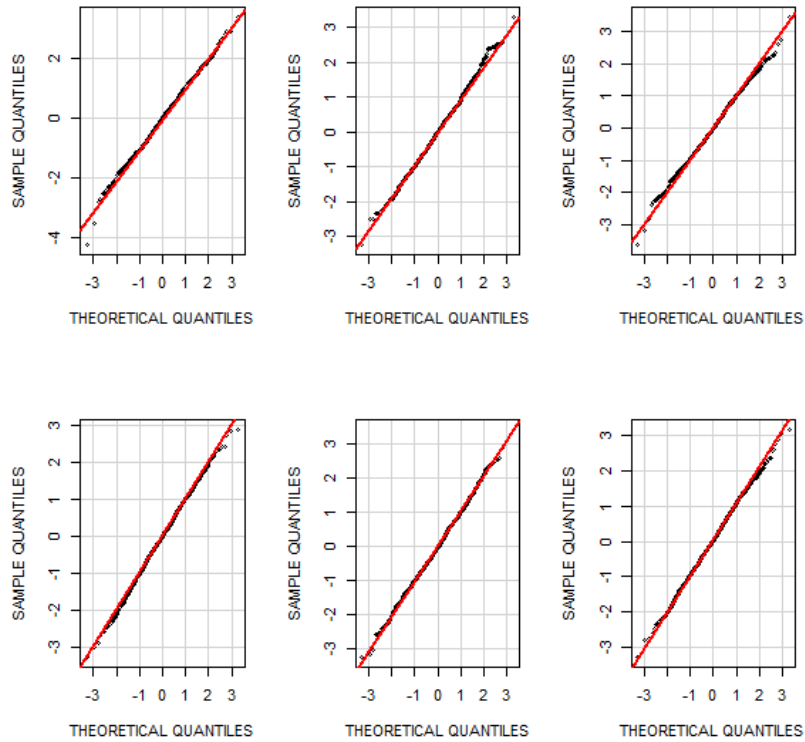| (n_1, n_2)	(μ_1, μ_2)			
	(1,1)	(1,2)	(2,3)	(2,5)
(50, 100)	0.943	0.943	0.944	0.942
(75, 100)	0.948	0.946	0.947	0.944
(100, 200)	0.950	0.948	0.950	0.949
(200, 300)	0.951	0.950	0.951	0.949
(500, 700)	0.954	0.950	0.956	0.951
(700, 1000)	0.957	0.958	0.957	0.958

Figure 1. the Q–Q plots against the standard normal distribution (Exponential populations)

First row:
Left: $(\mu_1, \mu_2) = (1,1)$ and $(n_1, n_2) = (50,100)$, Middle: $(\mu_1, \mu_2) = (1,2)$ and $(n_1, n_2) = (75,100)$, Right: $(\mu_1, \mu_2) = (1,2)$ and $(n_1, n_2) = (100,200)$.

Second row:
Left: $(\mu_1, \mu_2) = (2,3)$ and $(n_1, n_2) = (200,300)$, Middle: $(\mu_1, \mu_2) = (2,3)$ and $(n_1, n_2) = (500,700)$, Right: $(\mu_1, \mu_2) = (3,5)$ and $(n_1, n_2) = (700,1000)$.

Table 2. Shapiro-Wilk's normality test P-Value for the test statistic (Exponential populations)

(n_1, n_2)	(μ_1, μ_2)			
	(1,1)	(1,2)	(2,3)	(2,5)
(50, 100)	0.219	0.154	0.152	0.596
(75, 100)	0.253	0.293	0.337	0.598
(100, 200)	0.398	0.399	0.511	0.607
(200, 300)	0.403	0.540	0.720	0.649
(500, 700)	0.489	0.640	0.762	0.741
(700, 1000)	0.675	0.766	0.871	0.836

Example 4.2. Poisson populations

We assume X an d Y have Poisson distribution with mean μ_1 and μ_2, respectively. Table 3 presents the empirical coverage probabilities for different parameter setting.

Table 3. The empirical coverage probability (Poisson populations)

(n_1, n_2)	(μ_1, μ_2)			
	(1,1)	(1,2)	(2,3)	(2,5)
(50, 100)	0.942	0.942	0.943	0.948
(75, 100)	0.948	0.946	0.947	0.948
(100, 200)	0.953	0.949	0.949	0.949
(200, 300)	0.954	0.951	0.956	0.950
(500, 700)	0.959	0.959	0.956	0.953
(700, 1000)	0.960	0.959	0.957	0.954

As can be seen, in terms of the empirical coverage probability, the proposed method controls the type I error. Also Figure 2 and Table 4 show the Q–Q plots against the standard normal distribution and the results of Shapiro-Wilk's normality test for the test statistic in Eq. (3.4), respectively. It appears that our method performs well.

Table 4. Shapiro-Wilk's normality test P-Value for the test statistic (Poisson populations)

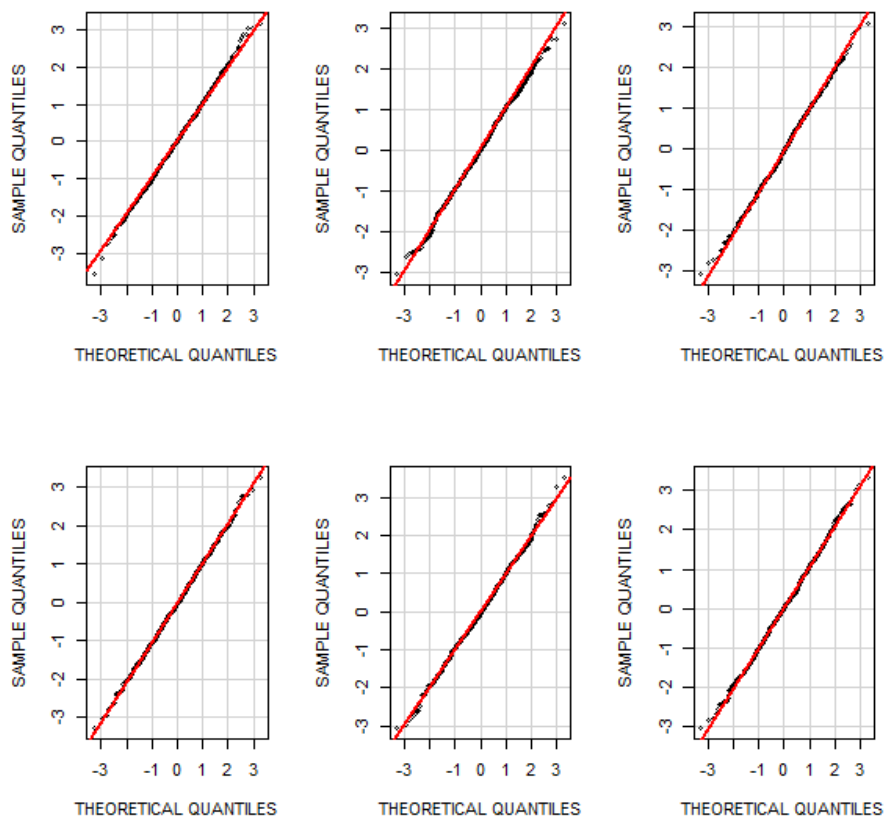| (n_1, n_2)	(μ_1, μ_2)			
	(1,1)	(1,2)	(2,3)	(2,5)
(50, 100)	0.156	0.179	0.263	0.185
(75, 100)	0.452	0.211	0.303	0.243
(100, 200)	0.553	0.701	0.346	0.403
(200, 300)	0.852	0.731	0.352	0.472
(500, 700)	0.885	0.893	0.538	0.739
(700, 1000)	0.887	0.916	0.861	0.928

Figure 2. the Q–Q plots against the standard normal distribution (Poisson populations)

First row:

Left: $(\mu_1, \mu_2) = (1,1)$ and $(n_1, n_2) = (50,100)$, Middle: $(\mu_1, \mu_2) = (1,2)$ and $(n_1, n_2) = (75,100)$, Right: $(\mu_1, \mu_2) = (1,2)$ and $(n_1, n_2) = (100,200)$.

Second row:

Left: $(\mu_1, \mu_2) = (2,3)$ and $(n_1, n_2) = (200,300)$, Middle: $(\mu_1, \mu_2) = (2,3)$ and $(n_1, n_2) = (500,700)$, Right: $(\mu_1, \mu_2) = (3,5)$ and $(n_1, n_2) = (700,1000)$.

Acknowledgement

References

[1] G. von der Emde, S. Schwarz, L. Gomez, R. Budelli, K.Grant, Electric fish measure distance in the dark. *Nature* **395** (1998) 890-894.

[2] D. J. Finney, *Statistical methods in biological assay (3 ed.)(Griffinn, London, 1978).*

[3] S. T. Hammett, P. G. Thompson, S. Bedingham, The dynamics of velocity adaptation in human vision. *Current Biology* **10** (2000) 1123-1126.

[4] M. Maes, A. Christophe, E. Bosmans, A. H. Lin, H. Neels, In humans, serum polyunsaturated fatty acid levels predict the response of prionflammatory cytokines to psychologic stress. *Biological Psychiatry* **47** (2000) 910-920.

[5] E. C. Fieller, The biological standardization of insulin. *Supplement to the Journal of the Royal Statistical Society* **7** (**1**) (1940) 1-64.

[6] E. C. Fieller, A fundamental formula in the statistics of biological assays and some applications. *Quarterly Journal of Pharmacy and Pharmacology* **17** (1944) 117-123.

[7] E. C. Fieller, Some problems in interval estimation. *Journal of the Royal Statistical Society (B)* **16** (**2**) (1954) 175-185.

[8] A. H. Briggs, C. Z. Mooney, D. E. Wonderling, Constructing confidence intervals for cost-effectiveness ratios: An evaluation of parametric and non-parametric techniques using Monte Carlo simulation. *Statistics in Medicine* **18** (**23**) (1999) 3245-3262.

[9] A. H. Briggs, B. J. O'Brien, G. Blackhouse, Thinking outside the box: Recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. *Annual Review of Public Health* **23** (2002) 377-401.

[10] A. C. Tamhane, B. R. Logan, Finding the maximum save dose level for heteroscedastic data. *Journal of Biopharmaceutical Statistics* **14** (2004) 843-856.

[11] D. Polsky, H. A. Glick, R. Willke, K. Schulman, Confidence intervals for cost-effectiveness ratios: A comparison of four methods. *Health Economics* **6** (**3**) (1997) 243-252.

[12] H. Wang, S. C. Chow, A practical approach for comparing means of two groups without equal variance assumption. *Statistics in Medicine* **21** (2002) 3137-3151.

[13] C. Cox, Fieller theorem, the likelihood and the delta-method. *Biometrics* **46** (**3**) (1990) 709-718.

[14] J. Gardiner, M. Huebner, J. Jetton, C. Bradley, On parametric confidence intervals for the cost-effectiveness ratio. *Biometrical Journal* **43** (**3**) (2001) 283-296.

[15] B. Efron, R. J. Tibshirani, *An introduction to the bootstrap* (Chapman & Hall, Boca Raton, 1993).

[16] M. A. Chaudhary, S. C. Stearns, Estimating confidence intervals for cost-effectiveness ratios: An example from a randomized trial. *Statistics in Medicine* **15** (1996) 1447-1458.

[17] K. B. Athreya, Bootstrap of the mean in the infinite variance case, *The Annals of Statistics* **15** (**2**) (1987) 724-731.

[18] K. Knight, On the bootstrap of the sample mean in the infinite variance case. *The Annals of Statistics* **17** (**3**) (1989) 1168-1175.

[19] J. T. G. Hwang, Fieller's problems and resampling techniques. *Statistica Sinica* **5** (1995) 161-171.

[20] H. Haghbin, M. R. Mahmoudi, Z. Shishebor, Large sample inference on the ratio of two independent binomial proportions, *Journal of Mathematical Extension* **5(1)** (2011) 87- 95.

[21] M. R. Mahmoudi, M. Mahmoodi, Inferrence on the ratio of variances of two independent populations. *Journal of Mathematical Extension* **7(2)** (2013) 83-91.

[22] M. R. Mahmoudi, M. Mahmoodi, Inferrence on the ratio of correlations of two independent populations. *Journal of Mathematical Extension* **7(4)** (2013) 71-82.

[23] M. R. Mahmouudi, M. Maleki, A. Pak, Testing the difference between two independent time series models, *Iranian Journal of Science and Technology (Sciences)* (2016).

[24] M. R. Mahmoudi, M. Mahmoudi, E. Nahavandi, Testing the difference between two independent regression models. *Communications in Statistics -- Theory and Methods* (2016), DOI: 10.1080/03610926.2014.960584.

[25] Thomas S. Ferguson, A course in large sample theory (Chapman & Hall, 1996).

[26] A. R. Willan, A. H. Briggs, *Statistical analysis of cost-effectiveness data*, (Chichester, England: John Wiley & Sons Ltd, 2006)