ASYMPTOTICS FOR LINEAR APPROXIMATIONS OF SMOOTH FUNCTIONS OF MEANS

Andrea Pallini

1. INTRODUCTION

The class of smooth functions of means contains several population characteristics and estimates, providing an important theoretical framework for statistical inference. The class of smooth functions is known to include the univariate variance, the difference of means, the ratio of means, the ratio of variances, the correlation coefficient, and the corresponding asymptotically pivotal versions. See Fuller (1976), chapter 5, Bhattacharya and Ghosh (1978), Hall (1992), chapter 2, Sen and Singer (1993), chapter 3, and Pallini (2000, 2002).

An accurate linear approximation for smooth functions of means is studied in Pallini (2002), with an error of order $O_p(n^{-2})$ in probability, as the sample size n diverges. Here, we study a higher-order version of the basic linear approximation studied in Pallini (2002), yielding an error in approximation of smaller order $O_p(n^{-3})$, as n diverges. We show that the asymptotic distribution of the basic and higher-order linear approximations is normal, as n diverges. We also study the corresponding sample versions of the basic and higher-order linear approximations, which can be used for approximating the sample estimate of a smooth function of means, with errors of order $O_p(n^{-5/2})$ and $O_p(n^{-7/2})$, respectively, as n diverges.

In particular, in section 2, we propose a higher-order version of the linear approximation of smooth functions of means proposed in Pallini (2002), by using a specific definition of difference of derivatives. We study a common asymptotics for the basic and its higher-order linear version, as n diverges. We show that these linear approximations are asymptotically normal, as n diverges. Finally, in section 3, we discuss the details of a simulation study on the ratio of means example, in order to see the computational issues and show the effectiveness of these linear approximations.

2. LINEAR APPROXIMATIONS

2.1. *Notation*

Let $\aleph = \{X_1, \ldots, X_n\}$ be a random sample of n independent and identically distributed (i.i.d.) observations drawn from a finite-dimensional q-variate random variable X, with distribution F and finite vector of means $\mu = E[X_1]$.

We denote by θ a real-valued population characteristic of interest. More precisely, θ is defined as $\theta = g(\mu)$, where $g: U \to \Re^1$ is a smooth function, $U \subseteq \Re^q$, fulfilling the conditions of Appendix 4.1. A natural estimator of θ is $\hat{\theta} = g(\overline{x})$, where $\overline{x} = n^{-1}\sum_{i=1}^{n} X_i$ is the vector of q sample means.

2.2. *The basic linear approximation*

In Pallini (2002), it is shown that

$$g(\overline{y} + \mu) - g(\mu) = \sum_{i=1}^{n} \{g(n^{-1} Y_i + \mu) - g(\mu)\} + O_p(n^{-2}) , \tag{1}$$

as $n \to \infty$, where $\overline{y} = n^{-1}\sum_{i=1}^{n} Y_i$, and $Y_i = X_i - \mu$, for every $i = 1, \ldots, n$. See also Appendix 4.2. The proof is detailed in Appendix 4.4.

Result (1) is obtained by Taylor expanding the smooth function $g(\overline{y} + \mu) - g(\mu)$ around the origin 0 of $\Re^q$, as the sum of the n Taylor expansions of the functions $g(n^{-1} Y_i + \mu) - g(\mu)$ around 0, where $i = 1, \ldots, n$.

The set of smooth functions $g(n^{-1} Y_i + \mu) - g(\mu)$, $i = 1, \ldots, n$, can be viewed as a sample of n well-defined i.i.d. random variables.

2.3. *The higher-order linear approximation*

For every $t_1 = 1, \ldots, q$, we denote by $g_{t_1}(u)$ the derivative of the smooth function $g(u)$, $g_{t_1}(u) = \partial g(u) / \partial u^{(t_1)}$, where $u \in U$. We define the quantity Γ as

$$\Gamma = \frac{1}{4} \left\{ \sum_{t_1=1}^{q} g_{t_1}(\overline{y} + \mu)\, \overline{y}^{(t_1)} - \sum_{i=1}^{n} \sum_{t_1=1}^{q} g_{t_1}(n^{-1} Y_i + \mu)\, n^{-1} Y_i^{(t_1)} \right\} , \tag{2}$$

where the derivatives are in the directions $\overline{y}$ and $n^{-1} Y_i$, at the points $\overline{y} + \mu$ and $n^{-1} Y_i + \mu$, respectively. The coefficient $1/4$ in (2) adjusts the difference of the derivatives for an exact correction of the term of order $O_p(n^{-2})$ in (1), as $n \to \infty$.

Adding Γ (given by (2)) to the basic linear approximation (1), we obtain the higher-order linear approximation given by

$$g(\overline{\mathrm{y}}+\mu)-g(\mu)=\sum_{i=1}^{n}\{g(n^{-1}\mathrm{Y}_{\mathrm{i}}+\mu)-g(\mu)\}+\Gamma+O_p(n^{-3})\,, \tag{3}$$

as $n\to\infty$. See Appendix 4.3 for the Taylor expansion of the derivative.

The order $O_p(n^{-3})$ in (3) improves on $O_p(n^{-2})$ in (1) by a good factor n^{-1}, as $n\to\infty$. See Appendix 4.5 for the proof. In Appendix 4.6, it is also seen that $\Gamma=O_p(n^{-1/2})$, as $n\to\infty$.

2.4. *Linear approximations of sample estimates*

The sample versions of the linear approximations (1) and (3) can naturally be obtained by substituting μ with $\overline{\mathrm{x}}$. These sample versions can be adapted to yield linear approximations of the sample estimate $g(\overline{\mathrm{x}})$ of the smooth function $g(\mu)$.

From (1) and (3), it follows that linear approximations of the estimates $\mathrm{g}(\overline{\mathrm{x}})$ can be defined by

$$g(\overline{\mathrm{x}})=n^{-1}\sum_{i=1}^{n} g(n^{-1}(\mathrm{X}_{\mathrm{i}}-\overline{\mathrm{x}})+\overline{\mathrm{x}})+O_p(n^{-5/2})\,, \tag{4}$$

as $n\to\infty$, and

$$g(\overline{\mathrm{x}})=n^{-1}\sum_{i=1}^{n} g(n^{-1}(\mathrm{X}_{\mathrm{i}}-\overline{\mathrm{x}})+\overline{\mathrm{x}})+n^{-1}2\hat{\Gamma}+O_p(n^{-7/2})\,, \tag{5}$$

as $n\to\infty$, where

$$\hat{\Gamma}=-\frac{1}{4}\sum_{i=1}^{n}\sum_{t_1=1}^{q} g_{t_1}(n^{-1}(\mathrm{X}_{\mathrm{i}}-\overline{\mathrm{x}})+\overline{\mathrm{x}})\,n^{-1}(\mathrm{X}_i^{(\mathrm{t}_1)}-\overline{\mathrm{x}}^{(\mathrm{t}_1)})\ . \tag{6}$$

The coefficient 2 in (5) adjusts $\hat{\Gamma}$ (given by (6)) for an exact correction of the term of order $O_p(n^{-5/2})$ in (4), as $n\to\infty$.

See Appendixes 4.7 and 4.8, for the proofs of the orders $O_p(n^{-5/2})$ and $O_p(n^{-7/2})$, as $n\to\infty$, in (4) and (5), respectively. In Appendix 4.9, it is also seen that $\hat{\Gamma}=O_p(n^{-1/2})$ and $n^{-1}\hat{\Gamma}=O_p(n^{-3/2})$, as $n\to\infty$.

In Appendixes 4.10 and 4.11, it is shown that the linear approximations (4) and

(5) of the estimates $g(\overline{\mathrm{x}})$ are asymptotically normal, as $n \to \infty$. In particular, it is seen that

$$n^{-1/2}\left\{\sum_{i=1}^{n} g(n^{-1}(\mathrm{X_i}-\overline{\mathrm{x}})+\overline{\mathrm{x}})-n\,g(\overline{\mathrm{x}})\right\} \xrightarrow{d} N(0,\sigma^2)\,, \tag{7}$$

$$n^{-1/2}\left\{\sum_{i=1}^{n} g(n^{-1}(\mathrm{X_i}-\overline{\mathrm{x}})+\overline{\mathrm{x}})+2\hat{\Gamma}-n\,g(\overline{\mathrm{x}})\right\} \xrightarrow{d} N(0,\sigma^2)\,, \tag{8}$$

as $n \to \infty$, where σ^2 is the asymptotic variance.

The asymptotic variance $n^{-1}\sigma^2$ in (7) and (8) can be estimated by

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \{g(n^{-1}(\mathrm{X_i}-\overline{\mathrm{x}})+\overline{\mathrm{x}})-g(\overline{\mathrm{x}})\}^2\,. \tag{9}$$

In Appendix 4.12, it is shown that $\hat{\sigma}^2 = n^{-1}\sigma^2 + O_p(n^{-3/2})$, as $n \to \infty$.

3. A SIMULATION STUDY

3.1. *The ratio of means example*

The random sample $\aleph$ consists of n i.i.d. bivariate observations $\mathrm{X_i} = (\mathrm{V_i^{(1)}}, \mathrm{V_i^{(2)}})^T$, where $q = 2$, and $i = 1,\ldots,n$. The random variable $\mathrm{V_1^{(2)}}$ ranges in a set of positive values. The population ratio of means is defined as

$$g(\mu) = \mu^{(1)}\,(\mu^{(2)})^{-1}\,.$$

The sample ratio of means $g(\overline{\mathrm{x}}) = \overline{\mathrm{x}}^{(1)}\,(\overline{\mathrm{x}}^{(2)})^{-1}$ is given by

$$g(\overline{\mathrm{y}}+\mu) = (\overline{\mathrm{y}}^{(1)}+\mu^{(1)})(\overline{\mathrm{y}}^{(2)}+\mu^{(2)})^{-1}\,.$$

The basic linear approximation (1) is defined by

$$\sum_{i=1}^{n}\{g(n^{-1}\mathrm{Y_i}+\mu)-g(\mu)\} = \sum_{i=1}^{n}\left\{\frac{n^{-1}\mathrm{Y}_i^{(1)}+\mu^{(1)}}{n^{-1}\mathrm{Y_i^{(2)}}+\mu^{(2)}}-\frac{\mu^{(1)}}{\mu^{(2)}}\right\}.$$

The quantity Γ (given by (2)) in the higher-order linear approximation (3) is defined by the derivative

$$\sum_{t_1=1}^{2} g_{t_1}(\overline{\mathrm{y}}+\mu)\;\overline{\mathrm{y}}^{(1)} = \frac{1}{\overline{\mathrm{y}}^{(2)}+\mu^{(2)}}\,\overline{\mathrm{y}}^{(1)} - \frac{\overline{\mathrm{y}}^{(1)}+\mu^{(1)}}{(\overline{\mathrm{y}}^{(2)}+\mu^{(2)})^2}\,\overline{\mathrm{y}}^{(2)}\,,$$

in the direction $\overline{\mathrm{y}}$ at the point $\overline{\mathrm{y}}+\mu$, and, for every $i=1,\ldots,n$, by the derivative

$$\sum_{t_1=1}^{2} g_{t_1}(n^{-1}\mathrm{Y}_i+\mu)\, n^{-1}\mathrm{Y}_{\mathrm{i}}^{(\mathrm{t}_1)} = \frac{1}{n^{-1}\mathrm{Y}_{\mathrm{i}}^{(2)}+\mu^{(2)}}\, n^{-1}\mathrm{Y}_{\mathrm{i}}^{(1)} - \frac{n^{-1}\mathrm{Y}_{\mathrm{i}}^{(1)}+\mu^{(1)}}{(n^{-1}\mathrm{Y}_{\mathrm{i}}^{(2)}+\mu^{(2)})^2}\, n^{-1}\mathrm{Y}_{\mathrm{i}}^{(2)},$$

in the direction $n^{-1}\mathrm{Y}_i$ at the point $n^{-1}\mathrm{Y}_i+\mu$.

Sample quantities in (4), (5) and (6) are given by the function

$$g(n^{-1}(\mathrm{X}_{\mathrm{i}}-\overline{\mathrm{x}})+\overline{\mathrm{x}}) = \frac{n^{-1}(\mathrm{X}_{\mathrm{i}}^{(1)}-\overline{\mathrm{x}}^{(1)})+\overline{\mathrm{x}}^{(1)}}{n^{-1}(\mathrm{X}_{\mathrm{i}}^{(2)}-\overline{\mathrm{x}}^{(2)})+\overline{\mathrm{x}}^{(2)}},$$

and, for every $i=1,\ldots,n$, by the derivative

$$\sum_{t_1=1}^{2} g_{t_1}(n^{-1}(\mathrm{X}_{\mathrm{i}}-\overline{\mathrm{x}})+\overline{\mathrm{x}})\, n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_1)}-\overline{\mathrm{x}}^{(\mathrm{t}_1)})$$

$$= \frac{1}{n^{-1}(\mathrm{X}_{\mathrm{i}}^{(2)}-\overline{\mathrm{x}}^{(2)})+\overline{\mathrm{x}}^{(2)}}\, n^{-1}(\mathrm{X}_{\mathrm{i}}^{(1)}-\overline{\mathrm{x}}^{(1)}) - \frac{n^{-1}(\mathrm{X}_{\mathrm{i}}^{(1)}-\overline{\mathrm{x}}^{(1)})+\overline{\mathrm{x}}^{(1)}}{(n^{-1}(\mathrm{X}_{\mathrm{i}}^{(2)}-\overline{\mathrm{x}}^{(2)})+\overline{\mathrm{x}}^{(2)})^2}\, n^{-1}(\mathrm{X}_{\mathrm{i}}^{(2)}-\overline{\mathrm{x}}^{(2)}).$$

3.2. *Computational details*

Figures 1 and 2 plot the differences between the sample ratio of means $g(\overline{x})$ and its linear approximations (4) and (5). Figures 3 and 4 compare the sample values of its linear approximations (4) and (5) with the corresponding quantiles of the standard normal distribution.

Simulated data were generated from a bivariate folded-normal distribution and a bivariate lognormal distribution. In particular, let $W_1=|N(0,1)|$, $W_2=|N(0,1)|$ and $W_3=|N(0,1)|$ be independent random variables, where $N(0,1)$ is the normal $N(0,1)$ random variable. The folded-normal variable $\mathrm{X}_1=(\mathrm{V}_1^{(1)},\mathrm{V}_1^{(2)})^T$ in Figure 1 is defined by $\mathrm{V}_1^{(1)}=W_1+W_3$ and $\mathrm{V}_1^{(2)}=W_2+W_3$, where the correlation coefficient of $\mathrm{V}_1^{(1)}$ and $\mathrm{V}_1^{(2)}$ is $\rho=0.5$. Let $W_1=N(0,1)$, $W_2=N(0,1)$ and $W_3=N(0,1)$ be independent random variables. The lognormal variable $\mathrm{X}_1=(\mathrm{V}_1^{(1)},\mathrm{V}_1^{(2)})^T$ in Figure 2 is defined by $\mathrm{V}_1^{(1)}=\exp((W_1+W_3)/\sqrt{2})$ and $\mathrm{V}_1^{(2)}=\exp((W_2+W_3)/\sqrt{2})$, where the correlation coefficient of $\mathrm{V}_1^{(1)}$ and $\mathrm{V}_1^{(2)}$ is $\rho=0.377541$.

3.3. *Empirical results*

Figures 1 and 2 show the performance of the sample linear approximations (4) and (5) for the ratio of means example. The simulated samples were generated with 40 different sizes n that range from $n=2$ to $n=41$. Linear approximations (4) and (5) may be equivalent in performance. It is important to see that the linear approximations (4) and (5) in Figure 1 can be regarded as nearly exact, for all sizes $n \geq 18$.

Figures 3 and 4 show the speed of convergence to normality of the linear approximations (4) and (5) for the ratio of means example, with small sample sizes n, $n=4$ and $n=7$. Results are preferable with data from the folded-normal distribution.

3.4. *Conclusions*

Sample linear approximations (4) and (5) are effective and very accurate. In any case, (5) may be more difficult than (4) to apply with some examples of smooth functions of means, since (5) demands the derivative of smooth functions.

Sample linear approximations (4) and (5) can be used for simplifying asymptotics in statistical inference for population smooth functions of means. Stochastic convergence of a sequence of smooth function of means can be studied by stochastic convergence of a sequence of sums of well-defined i.i.d. smooth functions that define the sample linear approximations (4) or (5).

Dipartimento di Scienze Statistiche "Paolo Fortunati" ANDREA PALLINI
Alma Mater Studiorum Università di Bologna

4. APPENDIX

4.1. *Assumptions*

We denote by $g_{t_1 \cdots t_k}(\mu)$ the derivative of order k of the smooth function $g(\mathrm{u}+\mu)$, $\mathrm{u} \in \mathrm{U}$, $g_{t_1 \ldots t_k}(\mu) = \partial^k g(\mathrm{u}+\mu) / \partial \mathrm{u}^{(t_1)} \ldots \partial \mathrm{u}^{(t_k)} \big|_{u=0}$, where $\mu = \mathrm{E}[\mathrm{Y}_1 + \mu]$. We let $\mu_{t_1 \cdots t_k} = \mathrm{E}[\mathrm{Y}_1^{(\mathrm{t}_1)} \cdots \mathrm{Y}_1^{(\mathrm{t}_k)}]$, and $\overline{\mathrm{y}} = n^{-1} \sum_{i=1}^{n} \mathrm{Y}_\mathrm{i}$.

Let $M_1 > 0$, $M_2 > 0$ and $M_3 > 0$ be finite constants. For every $k \leq s$, where $s \leq 5$, $g_{t_1 \cdots t_k}(\mu)$ exists and is bounded, $\left| g_{t_1 \cdots t_k}(\mu) \right| \leq M_1$. The sample smooth function $g(\overline{\mathrm{y}} + \mu)$ and the basic linear approximation $\sum_{i=1}^{n} \{g(n^{-1}\mathrm{Y}_\mathrm{i} + \mu) - g(\mu)\}$ are bounded,

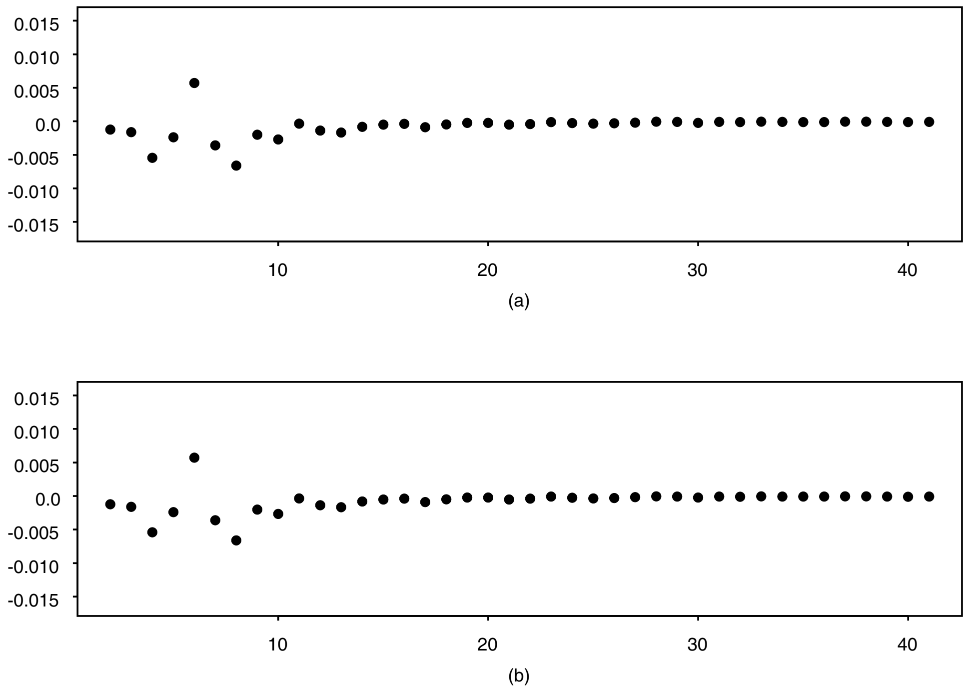

Figure 1 – Difference between the sample ratio of means $g(\overline{x})$ and its basic linear approximation given by (4) (panel (a)), and difference between $g(\overline{x})$ and its higher-order sample linear approximation given by (5) (panel (b)), for 40 different sample sizes n (horizontal axes) ranging from $n = 2$ to $n = 41$. Bivariate folded-normal observations from correlated ($\rho = 0.5$) marginals.

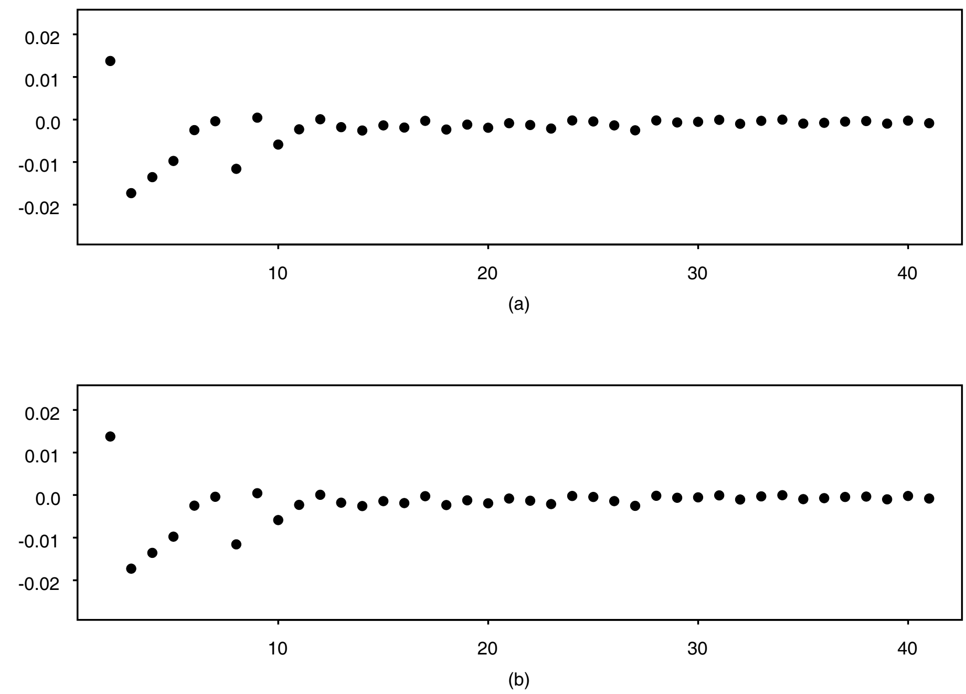

Figure 2 – Difference between the sample ratio of means $g(\overline{x})$ and its basic linear approximation given by (4) (panel (a)), and difference between $g(\overline{x})$ and its higher-order sample linear approximation given by (5) (panel (b)), for 40 different sample sizes n (horizontal axes) ranging from $n = 2$ to $n = 41$. Bivariate lognormal observations from correlated ($\rho = 0.377541$) marginals.

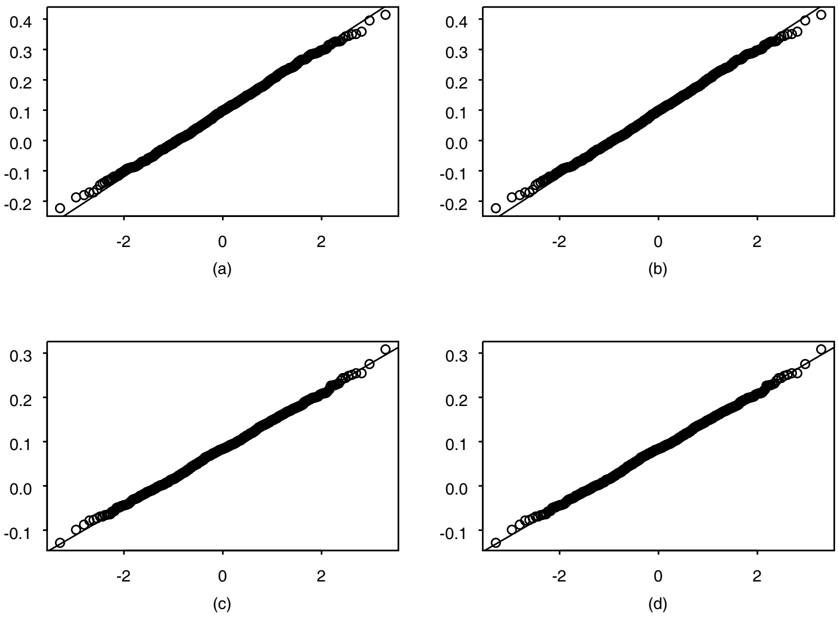

Figure 3 – Plot (x, y) of the quantiles of the standard normal distribution (x-axis) and the corresponding sample values of the linear approximation (4) (y-axis), sample size $n = 4$ (panel (a)) and $n = 7$ (panel (c)). Plot (x, y) of the quantiles of the standard normal distribution (x-axis) and the sample values of the higher-order linear approximation (5) (y-axis), sample size $n = 4$ (panel (b)) and $n = 7$ (panel (d)). The ratio of means example, with bivariate folded-normal observations from correlated ($\rho = 0.5$) marginals.

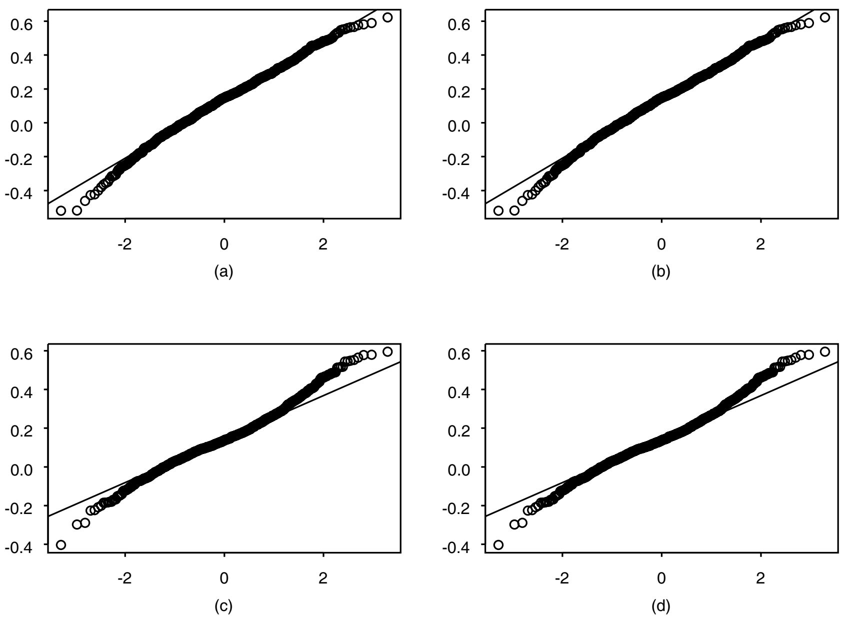

Figure 4 – Plot (x, y) of the quantiles of the standard normal distribution (x-axis) and the corresponding sample values of the linear approximation (4) (y-axis), sample size $n = 4$ (panel (a)) and $n = 7$ (panel (c)). Plot (x, y) of the quantiles of the standard normal distribution (x-axis) and the sample values of the higher-order linear approximation (5) (y-axis), sample size $n = 4$ (panel (b)) and $n = 7$ (panel (d)). The ratio of means example, with bivariate lognormal observations from correlated ($\rho = 0.377541$) marginals.

$$\left| \, g(\overline{\mathrm{y}} + \mu) \, \right| \le M_2 \,,$$

$$\left| \sum_{i=1}^{n} \{ g(n^{-1}\mathrm{Y_i} + \mu) - g(\mu) \} \right| \le M_3 \,.$$

4.2. *Exact linear approximations*

Following Appendix 4.1, if the derivative $g_{t_1 \cdots t_s}(\mu) = 0$, for all $s \ge 4$, then the basic linear approximation (1) is exact. If the derivative $g_{t_1 \cdots t_s}(\mu) = 0$, for all $s \ge 5$, then the higher-order linear approximation (3) is exact.

4.3. *Taylor expansion of the derivative*

For every $\mathrm{u} \in \mathrm{U}$, the Taylor expansion around $\mathrm{u} = 0$ of the derivative of the function $g(\mathrm{u} + \mu)$ is

$$\sum_{t_1=1}^{q} g_{t_1}(\mathrm{u} + \mu) \, \mathrm{u}^{(\mathrm{t}_1)} = \sum_{t_1=1}^{q} g_{t_1}(\mu) \, \mathrm{u}^{(\mathrm{t}_1)} + \sum_{k \ge 2} \frac{1}{(k-1)!} \sum_{t_1=1}^{q} \cdots \sum_{t_k=1}^{q} g_{t_1 \cdots t_k}(\mu) \, \mathrm{u}^{(\mathrm{t}_1)} \cdots \mathrm{u}^{(\mathrm{t}_k)} \,.$$

In particular, for every $\mathrm{u} \in \mathrm{U}$, we have that

$$\begin{aligned} \sum_{t_1=1}^{q} g_{t_1}(\mathrm{u} + \mu) \, \mathrm{u}^{(\mathrm{t}_1)} &= \sum_{t_1=1}^{q} g_{t_1}(\mathrm{u} + \mu) \, \mathrm{u}^{(\mathrm{t}_1)} \Big|_{\mathrm{u}=0} \\ &+ \left\{ \sum_{t_1=1}^{q} \sum_{t_2=1}^{q} g_{t_1 t_2}(\mathrm{u} + \mu) \, \mathrm{u}^{(\mathrm{t}_1)} \Big|_{\mathrm{u}=0} \mathrm{u}^{(\mathrm{t}_2)} \right. \\ &+ \left. \sum_{t_1=1}^{q} g_{t_1}(\mathrm{u} + \mu) \Big|_{\mathrm{u}=0} \mathrm{u}^{(\mathrm{t}_1)} \right\} \\ &+ \frac{1}{2} \left\{ \sum_{t_1=1}^{q} \sum_{t_2=1}^{q} \sum_{t_3=1}^{q} g_{t_1 t_2 t_3}(\mathrm{u} + \mu) \, \mathrm{u}^{(\mathrm{t}_1)} \Big|_{\mathrm{u}=0} \mathrm{u}^{(\mathrm{t}_2)} \mathrm{u}^{(\mathrm{t}_3)} \right. \\ &+ \left. 2 \sum_{t_1=1}^{\mathrm{q}} \sum_{t_2=1}^{q} g_{t_1 t_2}(\mathrm{u} + \mu) \Big|_{\mathrm{u}=0} \mathrm{u}^{(\mathrm{t}_1)} \mathrm{u}^{(\mathrm{t}_2)} \right\} \end{aligned}$$

$$+\frac{1}{6}\left\{\sum_{t_1=1}^{q}\sum_{t_2=1}^{q}\sum_{t_3=1}^{q}\sum_{t_4=1}^{q} \mathrm{g}_{\mathrm{t_1t_2t_3t_4}}(\mathrm{u}+\mu)\,\mathrm{u}^{(\mathrm{t_1})}\bigg|_{\mathrm{u}=0} \mathrm{u}^{(\mathrm{t_2})}\mathrm{u}^{(\mathrm{t_3})}\mathrm{u}^{(\mathrm{t_4})}\right.$$

$$\left.+3\sum_{t_1=1}^{q}\sum_{t_2=1}^{q}\sum_{t_3=1}^{q} g_{t_1t_2t_3}(\mathrm{u}+\mu)\bigg|_{\mathrm{u}=0} \mathrm{u}^{(\mathrm{t_1})}\mathrm{u}^{(\mathrm{t_2})}\mathrm{u}^{(\mathrm{t_3})}\right\}+\cdots.$$

4.4. *Proof of* $O_p(n^{-2})$ *in the basic linear approximation (1), as* $n \to \infty$

From the definition (1), by Taylor expanding the function $g(\mathrm{u}+\mu)$, where $\mathrm{u} \in U$, around $\mathrm{u}=0$, it follows that

$$g(\overline{\mathrm{y}}+\mu)-g(\mu)-\sum_{i=1}^{n}\{g(n^{-1}\mathrm{Y_i}+\mu)-g(\mu)\} = \sum_{k=1}^{4}\Delta_k + \mathrm{R},$$

where the quantities Δ_1, Δ_2, Δ_3 and Δ_4 are obtained as

$$\Delta_1 = \sum_{t_1=1}^{q} g_{t_1}(\mu)\,\overline{\mathrm{y}}^{(\mathrm{t_1})} - \sum_{i=1}^{n}\sum_{t_1=1}^{q} g_{t_1}(\mu)\,n^{-1}\mathrm{Y_i^{(t_1)}} = 0,$$

$$\Delta_2 = \frac{1}{2}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q} g_{t_1t_2}(\mu)\,\overline{\mathrm{y}}^{(\mathrm{t_1})}\overline{\mathrm{y}}^{(\mathrm{t_2})} - \frac{1}{2}\sum_{i=1}^{n}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q} g_{t_1t_2}(\mu)\,n^{-2}\mathrm{Y_i^{(t_1)}Y_i^{(t_2)}},$$

$$\Delta_3 = \frac{1}{6}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q}\sum_{t_3=1}^{q} g_{t_1t_2t_3}(\mu)\,\overline{\mathrm{y}}^{(\mathrm{t_1})}\overline{\mathrm{y}}^{(\mathrm{t_2})}\overline{\mathrm{y}}^{(\mathrm{t_3})}$$
$$-\frac{1}{6}\sum_{i=1}^{n}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q}\sum_{t_3=1}^{q} g_{t_1t_2t_3}(\mu)\,n^{-3}\mathrm{Y_i^{(t_1)}Y_i^{(t_2)}Y_i^{(t_3)}},$$

and

$$\Delta_4 = \frac{1}{24}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q}\sum_{t_3=1}^{q}\sum_{t_4=1}^{q} g_{t_1t_2t_3t_4}(\mu)\,\overline{\mathrm{y}}^{(\mathrm{t_1})}\overline{\mathrm{y}}^{(\mathrm{t_2})}\overline{\mathrm{y}}^{(\mathrm{t_3})}\overline{\mathrm{y}}^{(\mathrm{t_4})}$$
$$-\frac{1}{24}\sum_{i=1}^{n}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q}\sum_{t_3=1}^{q}\sum_{t_4=1}^{q} g_{t_1t_2t_3t_4}(\mu)\,n^{-4}\mathrm{Y_i^{(t_1)}Y_i^{(t_2)}Y_i^{(t_3)}Y_i^{(t_4)}},$$

respectively, and R is the remainder. Note that $\mathrm{E}[n^{-1}\mathrm{Y_1^{(t_1)}}]=0$, where $i=1,\ldots,n$, for every $t_1=1,\ldots,q$. For every $t_k=1,\ldots,q$, $k=1,2,3,4$, we have that

$$E[\overline{y}^{(t_1)}] = E\left[\sum_{i=1}^{n} n^{-1} Y_i^{(t_1)}\right] = 0 \ ,$$

$$E[\overline{y}^{(t_1)}\overline{y}^{(t_2)}] = E\left[\sum_{i=1}^{n} n^{-2} Y_i^{(t_1)} Y_i^{(t_2)}\right] = n^{-1}\mu_{t_1 t_2} \ ,$$

$$E[\overline{y}^{(t_1)}\overline{y}^{(t_2)}\overline{y}^{(t_3)}] = E\left[\sum_{i=1}^{n} n^{-3} Y_i^{(t_1)} Y_i^{(t_2)} Y_i^{(t_3)}\right] = n^{-2}\mu_{t_1 t_2 t_3} \ ,$$

$$E[\overline{y}^{(t_1)}\overline{y}^{(t_2)}\overline{y}^{(t_3)}\overline{y}^{(t_4)}] = n^{-3}\mu_{t_1 t_2 t_3 t_4} + O(n^{-2}) \ ,$$

$$E\left[\sum_{i=1}^{n} n^{-4}\, Y_i^{(t_1)} Y_i^{(t_2)} Y_i^{(t_3)} Y_i^{(t_4)}\right] = n^{-3}\mu_{t_1 t_2 t_3 t_4} \ ,$$

as $n \to \infty$. These results on multivariate moments show that $\Delta_1 = 0$ and $E[\Delta_2] = E[\Delta_3] = 0$. Then,

$$E\left[\left|\sum_{k=1}^{4} \Delta_k + R\right|\right] \geq \left|E[\Delta_4 + R]\right| = \lambda > 0.$$

There exists a finite constant $M_4 > 0$, such that $M_4\lambda = E[\,|\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4 + R|\,]$. For every $t > 0$, we also have that

$$\begin{aligned}
M_4\lambda &= \int_0^{+\infty} u\, dP\left(\left|\sum_{k=1}^{4} \Delta_k + R\right| \leq u \right) \\
&\geq \int_{t\lambda}^{+\infty} u\, dP\left(\left|\sum_{k=1}^{4} \Delta_k + R\right| \leq u \right) \\
&\geq t\lambda \int_{t\lambda}^{+\infty} dP\left(\left|\sum_{k=1}^{4} \Delta_k + R\right| \leq u \right) \\
&= t\lambda\, P\left(\left|\sum_{k=1}^{4} \Delta_k + R\right| > t\lambda \right).
\end{aligned}$$

Setting $t\lambda = \varepsilon$, where $\varepsilon > 0$, we obtain the Tchebychev inequality

$$P\left(\left| \sum_{k=1}^{4} \Delta_k + \mathrm{R} \right| > \varepsilon \right) \le \varepsilon^{-1} M_4 \lambda .$$

It finally follows that $\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4 + \mathrm{R} = O_p(n^{-2})$, as $n \to \infty$, because $M_4 = O(1)$ and $\lambda = O(n^{-2})$, as $n \to \infty$. Accordingly, it holds that $\mathrm{R} = O_p(n^{-3})$, as $n \to \infty$.

4.5. *Proof of* $O_p(n^{-3})$ *in the higher-order linear approximation (3), as* $n \to \infty$

Following Appendix 4.3, the quantity Δ_5 is defined as

$$\Delta_5 = \frac{1}{120} \sum_{t_1=1}^{q} \sum_{t_2=1}^{q} \sum_{t_3=1}^{q} \sum_{t_4=1}^{q} \sum_{t_5=1}^{q} g_{t_1 t_2 t_3 t_4 t_5}(\mu)\, \overline{\mathrm{y}}^{(\mathrm{t}_1)} \overline{\mathrm{y}}^{(\mathrm{t}_2)} \overline{\mathrm{y}}^{(\mathrm{t}_3)} \overline{\mathrm{y}}^{(\mathrm{t}_4)} \overline{\mathrm{y}}^{(\mathrm{t}_5)}$$
$$- \frac{1}{120} \sum_{i=1}^{n} \sum_{t_1=1}^{q} \sum_{t_2=1}^{q} \sum_{t_3=1}^{q} \sum_{t_4=1}^{q} \sum_{t_5=1}^{q} g_{t_1 t_2 t_3 t_4 t_5}(\mu)\, n^{-5} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_1)} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_2)} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_3)} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_4)} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_5)} .$$

From (3), it follows that

$$g(\overline{\mathrm{y}} + \mu) - g(\mu) - \sum_{i=1}^{n} \{ g(n^{-1}\mathrm{Y}_\mathrm{i} + \mu) - g(\mu) \} - \Gamma = \sum_{k=1}^{5} \Delta_k^* + \mathrm{R}^* ,$$

where Γ is given by (2), the quantities Δ_k^* are obtained as $\Delta_k^* = \Delta_k - k\Delta_k / 4$, $k = 1,2,3,4,5$, and R^* is the remainder. For every $t_k = 1,\ldots,q$, $k = 1,2,3,4,5$, we have that

$$\mathrm{E}[\overline{\mathrm{y}}^{(\mathrm{t}_1)} \overline{\mathrm{y}}^{(\mathrm{t}_2)} \overline{\mathrm{y}}^{(\mathrm{t}_3)} \overline{\mathrm{y}}^{(\mathrm{t}_4)} \overline{\mathrm{y}}^{(\mathrm{t}_5)}] = n^{-4} \mu_{t_1 t_2 t_3 t_4 t_5} + O(n^{-3}),$$

$$\mathrm{E}\left[\sum_{i=1}^{n} n^{-5} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_1)} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_2)} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_3)} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_4)} \mathrm{Y}_\mathrm{i}^{(\mathrm{t}_5)} \right] = n^{-4} \mu_{t_1 t_2 t_3 t_4 t_5} ,$$

as $n \to \infty$. Following Appendix 4.3, and these results on multivariate moments, we have that $\Delta_1^* = \Delta_4^* = 0$ and $\mathrm{E}[\Delta_2^*] = \mathrm{E}[\Delta_3^*] = 0$. We also have that $\left| \mathrm{E}[\Delta_5^* + R^*] \right| = \lambda^* > 0$, where $\lambda^* = O(n^{-3})$, as $n \to \infty$. Finally, it can be shown that $\Delta_1^* + \Delta_2^* + \Delta_3^* + \Delta_4^* + \Delta_5^* + \mathrm{R}^* = O_p(n^{-3})$, as $n \to \infty$, and (accordingly) $\mathrm{R}^* = O_p(n^{-4})$, as $n \to \infty$.

4.6. *Proof of* $\Gamma = O_p(n^{-1/2})$ *, as* $n \to \infty$ *, where* Γ *is given by (2)*

Following Appendix 4.3, we have that

$$\sum_{t_1=1}^{q} g_{t_1}(\overline{\mathrm{y}}+\mu)\,\overline{\mathrm{y}}^{(\mathrm{t}_1)} = \sum_{t_1=1}^{q} g_{t_1}(\mu)\,\overline{\mathrm{y}}^{(\mathrm{t}_1)} + O_p(n^{-1}),$$

as $n \to \infty$, and

$$\sum_{t_1=1}^{q} g_{t_1}(n^{-1}\mathrm{Y}_{\mathrm{i}}+\mu)\,n^{-1}\mathrm{Y}_{\mathrm{i}}^{(\mathrm{t}_1)} = \sum_{t_1=1}^{q} g_{t_1}(\mu)\,n^{-1}\mathrm{Y}_{\mathrm{i}}^{(\mathrm{t}_1)} + O_p(n^{-2}),$$

where $i=1,\ldots,n$, as $n \to \infty$, for every $t_1 = 1,\ldots,q$. It follows that

$$\Gamma = \frac{1}{4}\sum_{t_1=1}^{q} g_{t_1}(\mu)\,\overline{\mathrm{y}}^{(\mathrm{t}_1)} - \frac{1}{4}\sum_{i=1}^{n}\sum_{t_1=1}^{q} g_{t_1}(\mu)\,n^{-1}\mathrm{Y}_{\mathrm{i}}^{(\mathrm{t}_1)} + O_p(n^{-1}),$$

as $n \to \infty$. Finally, $\Gamma = O_p(n^{-1/2})$, as $n \to \infty$.

4.7. *Proof of* $O_p(n^{-5/2})$ *in the linear approximation (4), as* $n \to \infty$

For every $t_k = 1,\ldots,q$, $k = 1,2,3$, we have that

$$\sum_{i=1}^{n} n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_1)} - \overline{\mathrm{x}}^{(\mathrm{t}_1)}) = O_p(n^{-1/2}),$$

$$\sum_{i=1}^{n} n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_1)} - \overline{\mathrm{x}}^{(\mathrm{t}_1)})\,n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_2)} - \overline{\mathrm{x}}^{(\mathrm{t}_2)}) = O_p(n^{-3/2}),$$

$$\sum_{i=1}^{n} n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_1)} - \overline{\mathrm{x}}^{(\mathrm{t}_1)})\,n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_2)} - \overline{\mathrm{x}}^{(\mathrm{t}_2)})\,n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_3)} - \overline{\mathrm{x}}^{(\mathrm{t}_3)}) = O_p(n^{-5/2}),$$

as $n \to \infty$. Following Appendix 4.4, and substituting μ with $\overline{\mathrm{x}}$, we have that

$$\begin{aligned}\hat{\Delta}_1 &= -\sum_{i=1}^{n}\sum_{t_1=1}^{q} g_{t_1}(n^{-1}(\mathrm{X}_{\mathrm{i}} - \overline{\mathrm{x}}) + \overline{\mathrm{x}})\,n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_1)} - \overline{\mathrm{x}}^{(\mathrm{t}_1)}) \\ &= -\sum_{i=1}^{n}\sum_{t_1=1}^{q} \{g_{t_1}(\mu) + O_p(n^{-1/2})\}\,n^{-1}(\mathrm{X}_{\mathrm{i}}^{(\mathrm{t}_1)} - \overline{\mathrm{x}}^{(\mathrm{t}_1)}),\end{aligned}$$

$$\hat{\Delta}_2 = -\frac{1}{2}\sum_{i=1}^{n}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q} g_{t_1 t_2}(n^{-1}(\mathrm{X_i}-\overline{\mathrm{x}})+\overline{\mathrm{x}})\, n^{-1}(\mathrm{X_i^{(t_1)}}-\overline{\mathrm{x}}^{(\mathrm{t_1})})\, n^{-1}(\mathrm{X_i^{(t_2)}}-\overline{\mathrm{x}}^{(\mathrm{t_2})})$$
$$= -\frac{1}{2}\sum_{i=1}^{n}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q} \{g_{t_1 t_2}(\mu)+O_p(n^{-1/2})\}\, n^{-1}(\mathrm{X_i^{(t_1)}}-\overline{\mathrm{x}}^{(\mathrm{t_1})})\, n^{-1}(\mathrm{X_i^{(t_2)}}-\overline{\mathrm{x}}^{(\mathrm{t_2})}),$$

$$\hat{\Delta}_3 = -\frac{1}{6}\sum_{i=1}^{n}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q}\sum_{t_3=1}^{q} g_{t_1 t_2 t_3}(n^{-1}(\mathrm{X_i}-\overline{\mathrm{x}})+\overline{\mathrm{x}})$$
$$\cdot\, n^{-1}(\mathrm{X_i^{(t_1)}}-\overline{\mathrm{x}}^{(\mathrm{t_1})})\, n^{-1}(\mathrm{X_i^{(t_2)}}-\overline{\mathrm{x}}^{(\mathrm{t_2})})\, n^{-1}(\mathrm{X_i^{(t_3)}}-\overline{\mathrm{x}}^{(\mathrm{t_3})})$$
$$= -\frac{1}{6}\sum_{i=1}^{n}\sum_{t_1=1}^{q}\sum_{t_2=1}^{q}\sum_{t_3=1}^{q} \{g_{t_1 t_2 t_3}(\mu)+O_p(n^{-1/2})\}$$
$$\cdot\, n^{-1}(\mathrm{X_i^{(t_1)}}-\overline{\mathrm{x}}^{(\mathrm{t_1})})\, n^{-1}(\mathrm{X_i^{(t_2)}}-\overline{\mathrm{x}}^{(\mathrm{t_2})})\, n^{-1}(\mathrm{X_i^{(t_3)}}-\overline{\mathrm{x}}^{(\mathrm{t_3})}),$$

as $n \to \infty$. We also have that $\hat{\Delta}_1 = O_p(n^{-1})$, $\hat{\Delta}_2 = O_p(n^{-3/2})$ and $\hat{\Delta}_3 = O_p(n^{-5/2})$, as $n \to \infty$. It can be shown that $\hat{\Delta}_1 + \hat{\mathrm{R}} = O_p(n^{-1})$, where $\hat{\mathrm{R}}$ is the remainder, as $n \to \infty$, and (accordingly) $\hat{\mathrm{R}} = O_p(n^{-3/2})$, as $n \to \infty$. Then,

$$0 = \sum_{i=1}^{n} \{g(n^{-1}(\mathrm{X_i}-\overline{\mathrm{x}})+\overline{\mathrm{x}}) - g(\overline{\mathrm{x}})\} + O_p(n^{-3/2}),$$

as $n \to \infty$, and (4) is finally proved.

4.8. *Proof of $O_p(n^{-7/2})$ in the higher-order linear approximation (5), as $n \to \infty$*

Following Appendix 4.7, with $\hat{\Gamma}$ given by (6), we have that $\hat{\Delta}_1^* = \hat{\Delta}_2^* = 0$. It can be shown that $\hat{R}^* = O_p(n^{-5/2})$, where $\hat{R}^*$ is the remainder, as $n \to \infty$. Then, (5) is finally proved.

4.9. *Proof of $\hat{\Gamma} = O_p(n^{-1/2})$, as $n \to \infty$, where $\hat{\Gamma}$ is given by (6)*

Following Appendix 4.7, we have that

$$\hat{\Gamma} = -\frac{1}{4}\sum_{i=1}^{n}\sum_{t_1=1}^{q} \{g_{t_1}(\mu)+O_p(n^{-1/2})\}\, n^{-1}(\mathrm{X_i^{(t_1)}}-\overline{\mathrm{x}}^{(\mathrm{t_1})}),$$

as $n \to \infty$. Finally, we have that $\hat{\Gamma} = O_p(n^{-1/2})$ and $n^{-1}\hat{\Gamma} = O_p(n^{-3/2})$, as $n \to \infty$.

4.10. *Asymptotic normality of the linear approximation (4), as* $n \to \infty$

The variance $\mathrm{VAR}\left[\sum_{i=1}^{n} g(n^{-1}\mathrm{Y}_{\mathrm{i}} + \mu) - n\, g(\mu)\right]$ of the basic linear approximation (1) is

$$\sum_{i=1}^{n} \mathrm{VAR}\left[\sum_{t_1=1}^{q} g_{t_1}(\mu)\, n^{-1}\mathrm{Y}_{\mathrm{i}}^{(\mathrm{t}_1)} + O_p(n^{-2})\right] = n\,\{n^{-2}\sigma^2 + O(n^{-3})\}$$
$$= n^{-1}\sigma^2 + O(n^{-2}),$$

as $n \to \infty$, where

$$\sigma^2 = \sum_{t_1=1}^{q}\sum_{t_2=1}^{q} g_{t_1}(\mu) g_{t_2}(\mu)\, \mu_{t_1 t_2} .$$

The characteristic function $\phi_n(u)$ of

$$\sigma^{-1} n^{-1/2} \left\{ \sum_{i=1}^{n} g(n^{-1}\mathrm{Y}_{\mathrm{i}}^{(\mathrm{t}_1)} + \mu) - n\, g(\mu) \right\},$$

where $u \in \Re^1$, is

$$\phi_n(u) = \mathrm{E}\left[\exp\left(iu\,\sigma^{-1} n^{-1/2} \left\{ \sum_{i=1}^{n} g(n^{-1}\mathrm{Y}_{\mathrm{i}} + \mu) - n\, g(\mu) \right\}\right)\right]$$
$$= \mathrm{E}\left[1 + iu\,\sigma^{-1} n^{-1/2} \left\{ \sum_{i=1}^{n} g(n^{-1}\mathrm{Y}_{\mathrm{i}} + \mu) - n\, g(\mu) \right\}\right.$$
$$\left. + \frac{1}{2} i^2 u^2 \sigma^{-2} n^{-1} \left\{ \sum_{i=1}^{n} g(n^{-1}\mathrm{Y}_{\mathrm{i}} + \mu) - n\, g(\mu) \right\}^2 + \cdots \right]$$
$$= 1 - \frac{u^2}{2} + O(n^{-3/2}),$$

as $n \to \infty$. Actually, it holds that

$$\phi_n(u) = 1 - \frac{u^2}{2} + o(n^{-3/2}),$$

as $n \to \infty$. Then, we have that

$$\phi_n(u) \to \exp(-u^2/2),$$

as $n \to \infty$, which is the characteristic function of the normal $N(0,1)$ distribution. Following Appendixes 4.4 and 4.7, we finally have the asymptotic result (7).

4.11. *Asymptotic normality of the higher-order linear approximation (5), as* $n \to \infty$

In Appendix 4.4, it is seen that $\Delta_1 = 0$. The variance $\mathrm{VAR}\left[\sum_{i=1}^{n} g(n^{-1}\mathrm{Y}_\mathrm{i} + \mu) - n\, g(\mu) + \Gamma \right]$ of the higher-order linear approximation (3) then is

$$\sum_{i=1}^{n} VAR\left[\sum_{t_1=1}^{q} g_{t_1}(\mu)\, n^{-1}\mathrm{Y}_\mathrm{i}^{(\mathrm{t}_1)} + O_p(n^{-2}) \right] = n^{-1}\sigma^2 + O(n^{-2}),$$

as $n \to \infty$, where σ^2 is defined in Appendix 4.10.

The characteristic function $\phi_n^*(u)$ of

$$\sigma^{-1} n^{-1/2} \left\{ \sum_{i=1}^{n} g(n^{-1}\mathrm{Y}_\mathrm{i} + \mu) - n\, g(\mu) + \Gamma \right\},$$

where $u \in \Re^1$, is

$$\phi_n^*(u) = 1 - \frac{u^2}{2} + o(n^{-5/2}),$$

as $n \to \infty$. Then, we have that

$$\phi_n^*(u) \to \exp(-u^2/2),$$

as $n \to \infty$, which is the characteristic function of the normal $N(0,1)$ distribution. Following Appendixes 4.5, 4.8 and 4.9, we finally have the asymptotic result (8).

4.12. *Proof of* $\hat{\sigma}^2 = n^{-1}\sigma^2 + O_p(n^{-3/2})$ *, as* $n \to \infty$ *, where* $\hat{\sigma}^2$ *is given by (9)*

From the definition (9), we have that

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \left\{ \sum_{t_1=1}^{q} g_{t_1}(\mu)\, n^{-1}\mathrm{Y}_\mathrm{i}^{(\mathrm{t}_1)} + O_p(n^{-3/2}) \right\}^2,$$

as $n \to \infty$. The asymptotic variance σ^2 is defined in Appendix 4.10. Then, it finally follows that $\hat{\sigma}^2 = n^{-1}\sigma^2 + O_p(n^{-3/2})$, as $n \to \infty$.

REFERENCES

R.A. BECKER, J.M. CHAMBERS, A.R. WILKS (1988), *The New S Language*, Wadsworth and Brooks/Cole, Pacific Grove, California.

R.N. BHATTACHARYA, J.K. GHOSH (1978), *On the validity of the formal Edgeworth expansion*, "Annals of Statistics", 6, 434-451.

W.A. FULLER (1976), *Introduction to Statistical Time Series*, Wiley and Sons, New York.

P. HALL (1992), *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York.

A. PALLINI (2000), *Efficient bootstrap estimation of distribution functions,* "Metron", 58 (1-2), 81-95.

A. PALLINI (2002), *On a linear method in bootstrap confidence intervals*, "Statistica", 62, 5-25.

P.K. SEN, J.M. SINGER (1993), *Large Sample Methods in Statistics,* Chapman and Hall, New York.

RIASSUNTO

Analisi asintotica per approssimazioni lineari di medie

Viene definita e studiata una versione di ordine più accurato della linearizzazione di funzioni regolari di medie proposta in Pallini (2002). Viene dimostrato come questa versione migliori l'errore di approssimazione di ordine $O_p(n^{-2})$ in probabilità, al divergere della numerosità campionaria n, producendo un errore più piccolo di ordine $O_p(n^{-3})$, al divergere di n. Viene dimostrata la normalità asintotica di entrambe le linearizzazioni al divergere di n. Vengono presentati i risultati empirici di uno studio di simulazione sull'esempio del rapporto di due medie.

SUMMARY

Asymptotics for linear approximations of smooth functions of means

A higher-order version of the linear approximation of smooth functions of means proposed in Pallini (2002) is defined and studied. This version is shown to improve over the error of order $O_p(n^{-2})$ in probability, as the sample size n diverges, yielding a smaller error of order $O_p(n^{-3})$, as n diverges. Both linear approximations are shown to have a normal distribution, as n diverges. Empirical results of a simulation study on the ratio of means example are presented.