

Asymptotic expansions for conditional moments of Bernoulli trials

Strasser, Helmut

Published in:
Statistics and Risk Modeling

DOI:
[10.1524/strm.2012.1124](https://doi.org/10.1524/strm.2012.1124)

Published: 01/12/2012

Document Version
Peer reviewed version

[Link to publication](#)

Citation for pulished version (APA):
Strasser, H. (2012). Asymptotic expansions for conditional moments of Bernoulli trials. *Statistics and Risk Modeling*, 21(4), 327 - 343. <https://doi.org/10.1524/strm.2012.1124>

Asymptotic expansions for conditional moments of Bernoulli trials

Helmut Strasser*

September 24, 2012

Abstract

In this paper we study conditional distributions of independent, but not identically distributed Bernoulli random variables. The conditioning variable is the sum of the Bernoulli variables. We obtain Edgeworth expansions for the conditional expectations and the conditional variances and covariances. The results are of basic interest for several applications, e.g. for the study of conditional maximum likelihood estimation in Rasch models with many item parameters.

1 Introduction

Let X_1, X_2, \dots, X_n independent random variables and let $S_n := X_1 + X_2 + \dots + X_n$ be their sum. If the random variables are identically distributed it is well-known how to study the conditional distribution $P(\cdot | S_n)$. Given the sum, the variables are exchangeable. Conditional expectations and covariances are then tractable by simple explicit expressions. For the process of the partial sums there is a conditional invariance principle whose limit is a Brownian bridge (see Billingsley [2]).

The situation is much more complicated if the random variables are not identically distributed. The first problem is to find expressions for the conditional expectations and covariances. If this is solved then the way is open for studying asymptotics of the conditional distributions.

The present paper is concerned with those questions for the case of random variables with values 0 and 1 and probabilities $P(X_i = 1) = p_i$. The numerical calculation of conditional probabilities $P(\cdot | S_n)$ is computer intensive and does not answer any structural

*Institute for Statistics and Mathematics, WU, Vienna

question. It is much more informative to use asymptotics like a local limit theorem and the corresponding asymptotic expansions. This purpose is supported by a wealth of excellent literature on the theory of asymptotic expansions related to central and local limit theorems, e.g. Bhattacharya and Rao, [1]. In the present paper we apply this theory to obtain asymptotic expansions for the conditional expectations and covariances.

The main results are Theorems 2.1 , 2.3 and 2.4 . These results are basic for further research on the conditional distributions. For reasons of space such applications will be treated in subsequent papers. However, in the following we briefly describe two applications in order to illustrate the power of the results of the present paper.

An important application is the proof of an invariance principle for the partial sum process under the conditional distribution given the sum. The corresponding assertion for i.i.d. variables is well-known since 1968 (Billingsley [2]). A simplified proof can be given by martingale methods and can be found in Hoffmann [3]. These martingale methods can be extended to not identically distributed variables as soon as there are sufficiently precise expressions for conditional expectations and covariances.

Another application is concerned with the so-called Rasch model of psychometrics. In this case the conditional distributions $P(\cdot|S_n)$ are used for parameter estimation. The results of the present paper make it possible to clarify the asymptotic covariance structure of conditional maximum likelihood estimates when the number of item parameters is large. For details see Strasser [4].

In this paper only Bernoulli variables are considered. In this way it is shown that and how the proving strategy works in principle. The strategy obviously can be extended to more general random variables using the theory presented in [1]. However, then the situation is more complicated and more complex assumptions have to be imposed.

In the proofs we present those arguments and steps which are indispensable for understanding the logical structure and the main calculations. Boring routine calculations (like multiplying polynomials and sorting their terms) are left to the reader under the label of "simple calculations". These steps have been performed by the author with the aid of machine support (software package Maxima). A detailed documentation is contained in Strasser [5].

2 The main results

All results are proved under the assumption that the sequence of probabilities (p_i) is contained in a closed subinterval of $(0, 1)$.

Let us introduce some notation. Let $v_i := p_i(1 - p_i)$, $\sigma_n^2 := \sum_{i=1}^n v_i$ and $\bar{\sigma}_n^2 = \sigma_n^2/n$. In

general, the average of any vector $\mathbf{a} = (a_1, \dots, a_n)$ is denoted by \bar{a} . Moreover, we denote

$$\tau_{ni} := 2 \left(p_i - \sum_{j=1}^n \frac{v_j}{\sigma_n^2} p_j \right).$$

For identically distributed random variables this number is zero.

For denoting the order of convergence we use the familiar Landau notation $O(n^{-k})$ and $o(n^{-k})$. If we talk about convergence of functions then $O_c(n^{-k})$ and $o_c(n^{-k})$ mean uniform convergence of compact subsets.

Let $Z_n := (S_n - n\bar{p})/\sigma_n$. These random variables are standardized, asymptotically normally distributed and have uniformly bounded moments of any order. Clearly, the sequence (Z_n) is asymptotically normally distributed and stochastically bounded. More precise information in this direction is given in section 4.

The first result is Theorem 2.1 which is concerned with conditional expectations $E(X_i|S_n)$. As a preparation let us quote the corresponding assertion for i.i.d. variables. In this case we have

$$E(X_i|S_n) = \frac{S_n}{n} = p + \frac{1}{\sqrt{n}} \frac{v}{\sigma_n} Z_n.$$

It is easy to see that this conditional expectation is equal to the best linear predictor (linear regression). Theorem 2.1 shows that this structure of the conditional expectation remains true in general with precision $O(n^{-1})$. Moreover, we identify the correction term of order n^{-1} .

2.1 THEOREM. *The expansion*

$$E(X_i|S_n) = p_i + \frac{1}{\sqrt{n}} \frac{v_i}{\sigma_n} Z_n - \frac{1}{n} \frac{v_i \tau_{ni}}{2\sigma_n^2} (Z_n^2 - 1) + r_{ni}(Z_n)$$

holds with $E(|r_{ni}(Z_n)|) = O(n^{-3/2})$ and $r_{ni}(z) = O_c(n^{-3/2})$.

Proof: This is a consequence of Theorem 5.4 for $k = 1$, where the expansion is used up to the term of order n^{-1} . The explicit expression for the coefficients is obtained from Lemma 6.1 by easy calculations and expanding the abbreviations of (9). \square

The second result is Theorem 2.3 which will be concerned with conditional variances and covariances. First, Lemma 2.2 contains the expansions of the second moments $E(X_i X_j | S_n)$. As a preparation let us again quote the corresponding assertion for i.i.d. variables. For $i \neq j$ and apart from a remainder term of order $n^{-3/2}$ we have in this case

$$E(X_i X_j | S_n) = \frac{S_n(S_n - 1)}{n(n - 1)} \approx p^2 + \frac{1}{\sqrt{n}} \frac{2pv}{\sigma_n} Z_n + \frac{1}{n} \frac{v^2}{\sigma_n^2} (Z_n^2 - 1). \quad (1)$$

Lemma 2.2 shows how things change if the random variables are not identically distributed.

2.2 LEMMA. *If $i \neq j$ the expansion*

$$E(X_i X_j | S_n) = p_i p_j + \frac{1}{\sqrt{n}} \frac{p_i v_j + p_j v_i}{\bar{\sigma}_n} Z_n + \frac{1}{n} \frac{v_i v_j}{\bar{\sigma}_n^2} (Z_n^2 - 1) - \frac{1}{n} \frac{p_i v_j \tau_{nj} + p_j v_i \tau_{ni}}{2\bar{\sigma}_n^2} (Z_n^2 - 1) + r_{nij}(Z_n)$$

holds with $E(|r_{nij}(Z_n)|) = O(n^{-3/2})$ and $r_{nij}(z) = O_c(n^{-3/2})$.

Proof: This is also a consequence of Theorem 5.4 for $k = 2$, where the expansion is used up to the term of order n^{-1} . The explicit expression for the coefficients is again obtained from Lemma 6.1 by easy calculations and expanding the abbreviations of (9). \square

It is not difficult to obtain from Theorem 2.1 and Lemma 2.2 the corresponding expansions of the conditional covariances. It is interesting to compare them with the i.i.d. case where the covariances are

$$E(X_i X_j | S_n) - E(X_i | S_n) E(X_j | S_n) = \frac{S_n(S_n - 1)}{n(n - 1)} \left(\delta_{ij} - \frac{1}{n} \right).$$

In view of (1) we see that only the diagonal entries ($i = j$) contain terms of order n^0 and $n^{-1/2}$. If $i \neq j$ all terms are equal and of the order n^{-1} . Hence, the structure of the following assertion is not surprising.

2.3 THEOREM. *For conditional variances the expansion*

$$E(X_i^2 | S_n) - E(X_i | S_n)^2 = v_i - \frac{1}{n} \frac{v_i^2}{\bar{\sigma}_n^2} + \frac{1}{\sqrt{n}} \frac{v_i(1 - 2p_i)}{\bar{\sigma}_n} Z_n - \frac{1}{n} \frac{v_i^2}{\bar{\sigma}_n^2} (Z_n^2 - 1) - \frac{1}{n} \frac{v_i(1 - 2p_i)\tau_{ni}}{2\bar{\sigma}_n^2} (Z_n^2 - 1) + r_{ni}(Z_n)$$

holds with $E(|r_{ni}(Z_n)|) = O(n^{-3/2})$ and $r_{ni}(z) = O_c(n^{-3/2})$.

For conditional covariances with $i \neq j$ the expansion

$$E(X_i X_j | S_n) - E(X_i | S_n) E(X_j | S_n) = -\frac{1}{n} \frac{v_i v_j}{\bar{\sigma}_n^2} + r_{nij}(Z_n)$$

holds with $E(|r_{nij}(Z_n)|) = O(n^{-3/2})$ and $r_{nij}(z) = O_c(n^{-3/2})$.

The third result is Theorem 2.4 and is concerned with the expectations of the conditional variances and covariances. For i.i.d. variables we have

$$E\left(E(X_i X_j | S_n) - E(X_i | S_n)E(X_j | S_n)\right) = \frac{v^2}{\sigma_n^2} \left(\delta_{ij} - \frac{1}{n}\right).$$

Expansions of order $O(n^{-3/2})$ follow immediately from Theorem 7. However, we identify the terms of order n^{-2} .

2.4 THEOREM. *The expectations of the conditional variances satisfy*

$$E\left(E(X_i^2 | S_n) - E(X_i | S_n)^2\right) = v_i - \frac{1}{n} \frac{v_i^2}{\sigma_n^2} + \frac{1}{n^2} \frac{v_i^2 \tau_{ni}^2}{2\sigma_n^4} + O(n^{-5/2}).$$

The expectations of the conditional covariances for $i \neq j$ satisfy

$$E\left(E(X_i X_j | S_n) - E(X_i | S_n)E(X_j | S_n)\right) = -\frac{1}{n} \frac{v_i v_j}{\sigma_n^2} + \frac{1}{n^2} \frac{v_i v_j \tau_{ni} \tau_{nj}}{2\sigma_n^4} + O(n^{-5/2}).$$

For the proof of Theorem 2.4 we refer to section 5.

3 Numerical experiments

The following numerical results are taken from Strasser [6] where further details on algorithms and coding are presented.

The first experiment is concerned with $n = 30$ Bernoulli variables whose parameters p_i are chosen randomly from a uniform distribution on $(0, 1)$. Table 1 shows the statistical measures of the differences between the lhs and the rhs of Theorem 2.1 .

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-0.340600	-0.095110	-0.002954	-0.051520	0.001969	0.079250

Table 1: Approximation of $E(X_1 | S_n = s)$, $s = 0, 1, \dots, n$

Table 2 shows the statistical measures of the differences between the lhs and the rhs of Lemma 2.2 .

Both examples indicate that the error distributions have heavy tails. This is compatible with theory which says that the approximations are uniform for values $s = 0, 1, \dots, n$ of

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-0.2740000	-0.0644400	-0.0009374	-0.0376500	0.0106400	0.1258000

Table 2: Approximation of $E(X_1 X_2 | S_n = s)$, $s = 0, 1, \dots, n$

S_n only as long as (Z_n) is uniformly bounded. In view of the central limit theorem the sequence of random variables (Z_n) is uniformly bounded with probabilities arbitrarily near to one, i.e.

$$\sup_n P(|Z_n| \geq a_\epsilon) < \epsilon \text{ for every } \epsilon > 0 \text{ and suitable } a_\epsilon < \infty.$$

In this sense the approximations are valid with probabilities arbitrary close to one. In order to check the validity of the theoretical results we have to consider regions of $s = 0, 1, \dots, n$ where (Z_n) remains bounded with large probability. In the following we calculate the maximal absolute errors of the approximation polynomials on the range where $|Z_n| < N_{0.995}$ (where N_α denotes the α -quantile of the normal distribution).

Now we study a sequence of approximations based on a randomly chosen vector \mathbf{p} of length $n = 300$. We compute the maximal absolute error between the exact values and the approximate values for the sequence of vectors $\mathbf{p}_k = (p_1, \dots, p_k)$, $k = 1, 2, \dots, n$.

When we consider the \log_{10} - \log_{10} plots in Figure 1 of the maximal absolute errors with respect to the length of \mathbf{p} then we observe a linear pattern which indicates a power function relation. The \log_{10} - \log_{10} plot also displays the slope of the regression line. Recall that the theoretical results of Theorem 2.1 and Lemma 2.2 predict a slope of 1.5. The numerical results support the quality of the theoretical error bounds.

The result of Theorem 2.3 on the approximation of the conditional covariance matrix is a direct consequence of the preceding results and need not be illustrated any further.

Things are becoming much more interesting when we pass to the expectation of the conditional covariance matrix, i.e. to the assertion of Theorem 2.4. These approximations are of a considerably higher order and they do not require any truncation of tails. Moreover it is a matter of considerable interest that the approximation of order $1/n$ holds already for small values of n with great precision.

We randomly choose 100 vectors \mathbf{p} of length $n = 20$ and study the maximum of the absolute deviations between exact values of the expected conditional covariances and their approximations. We do that both for the approximations of order $1/n$ and of order $1/n^2$. Tables 3 and 4 show the statistical measures of those deviations. The results are more than satisfactory.

Finally we study a sequence of approximations based on a randomly chosen vector \mathbf{p} of length $n = 50$. We compute the maximal absolute error between the exact values and the

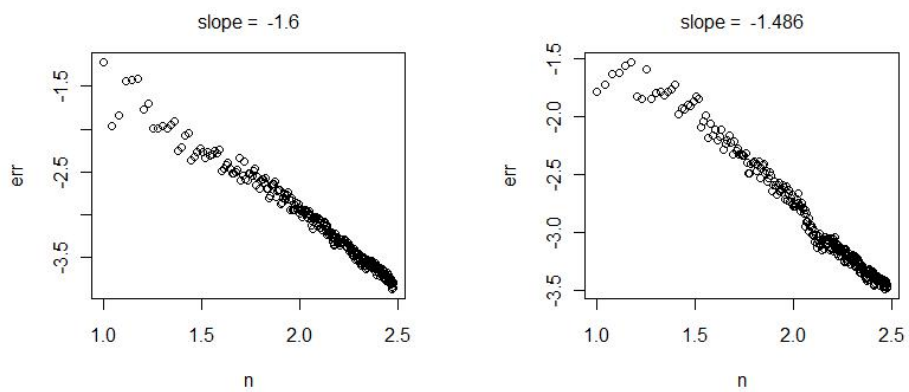


Figure 1: Approximation of $E(X_1|S_n)$ and $E(X_1X_2|S_n)$

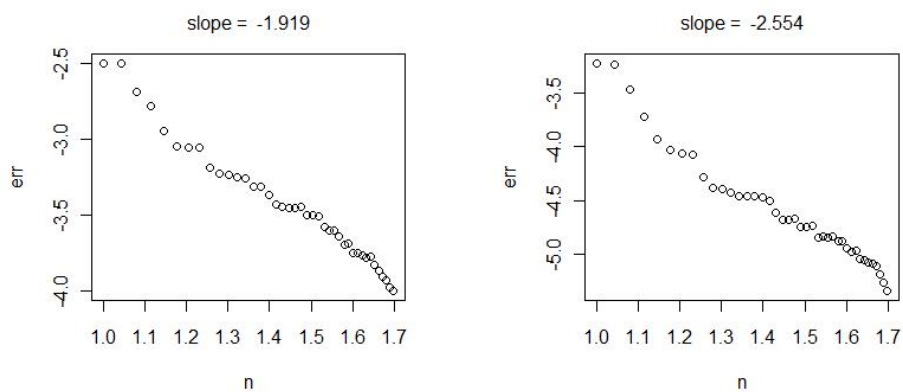


Figure 2: Approximation of Theorem 2.4

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.0003435	0.0004523	0.0005354	0.0005738	0.0006347	0.0015420

Table 3: Approximation in Theorem 2.4 of order $1/n$

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2.133e-05	3.531e-05	4.502e-05	5.468e-05	6.122e-05	2.304e-04

Table 4: Approximation in Theorem 2.4 of order $1/n^2$

approximate values for all vectors $\mathbf{p}_k = (p_1, \dots, p_k)$, $k = 1, 2, \dots, n$. When we consider the \log_{10} - \log_{10} plots in Figure 2 then we observe again a linear pattern which indicates a power function relation. The slope of the log-log plot estimates the exponent of the order of decrease. It supports the theoretical results perfectly.

4 Edgeworth expansions

The results of the present paper are based on theory of asymptotic expansions which is presented in the monograph [1] by R. N. Bhattacharya and R. R. Rao. In this section we are going to explain basic notation and facts. We continue using assumptions and notations of section 2 of the present paper.

Our results are based on Theorem 22.1 of Bhattacharya-Rao [1]. Theorem 22.1 from Bhattacharya-Rao [1] provides a so-called Edgeworth-expansion of probabilities

$$P(S_n = s_n) = P(Z_n = z_n) =: f_n(z_n), \quad (2)$$

where $z_n := \frac{s_n - n\bar{p}}{\sigma_n}$. See chapter 6, section 6, of Bhattacharya-Rao [1] for a nice description of the Fourier-transform machinery which leads to Edgeworth-expansions.

In order to give the terms of the expansion we need further notation. Let χ_{ki} be the cumulant of X_i of order k , i.e.

$$\chi_{ki} := \frac{d^k}{dt^k} (\log(1 - p_i + p_i \exp(t)))|_{t=0}.$$

We also need abbreviations for average cumulants

$$\bar{\chi}_{nk} := \frac{1}{n} \sum_{i=1}^n \chi_{ki}, \quad \bar{\sigma}_n^2 := \bar{\chi}_{n2}.$$

and for the coefficients

$$w_{nk} := \frac{\bar{\chi}_{nk}}{\sigma_n^k}.$$

In the present paper only expansion up to order $r = 2$ are considered. The corresponding so-called Edgeworth-polynomials are

$$P_{n1}(z) := \frac{w_{n3}}{3!} H_3(z)$$

$$P_{n2}(z) := \frac{w_{n4}}{4!} H_4(z) + \frac{w_{n3}^2}{2!(3!)^2} H_6(z)$$

where H_k denote Hermite polynomials

$$H_3(x) := x^3 - 3x,$$

$$H_4(x) := x^4 - 6x^2 + 3,$$

$$H_6(x) := x^6 - 15x^4 + 45x^2 - 15.$$

Theorem 22.1 in Bhattacharya-Rao [1] is formulated for a sequence of i.i.d. random variables. For the results of the present paper, however, the validity of Theorem 22.1 for independent but not identically distributed random variables is essential. This is discussed in Bhattacharya-Rao [1] on pages 240/241 where also the corresponding modifications of assumptions and facts are explained.

Now we present that modification of Theorem 22.1 from Bhattacharya-Rao [1] which holds for a sequence of independent but not identically distributed Bernoulli variables.

4.1 THEOREM. *Assume that the sequence (p_i) is contained in a closed subinterval of $(0, 1)$. Then*

$$g_{nr}(z) := \frac{1}{\sigma_n} \phi(z) \left(1 + \sum_{j=1}^r \frac{1}{n^{j/2}} P_{nj}(z) \right)$$

satisfies

$$\sup_{z_n} |f_n(z_n) - g_{nr}(z_n)| = O(n^{-(r+1)/2}) \quad (3)$$

and

$$\sum_{z_n} |f_n(z_n) - g_{nr}(z_n)| = O(n^{-r/2}). \quad (4)$$

It should be noted that by (3) the distances are uniformly bounded which implies that (4) is valid for all powers with exponents ≥ 1 , too.

Theorem 4.1 is formulated and proved in Bhattacharya-Rao [1]. For our purposes the order of approximation in (3) can be modified. Let us define a stochastic order of approximation.

4.2 DEFINITION. Let us say that a sequence $(f_n(z))$ of functions satisfies $f_n(Z_n) = O_Z(n^{-k})$ if there are constants C and m such that

$$|f_n(Z_n)| \leq \frac{C}{n^k} |Z_n|^m, \quad n \in \mathbb{N}.$$

If $f_n(Z_n) = O_Z(n^{-k})$ then it follows from Lemma 6.1 :

(1) The moments satisfy

$$E(|f_n(Z_n)|^p) = O(n^{-kp}) \text{ if } p > 0. \quad (5)$$

(2) For any $r \in \mathbb{N}$ there are constants $a > 0$ and $b > 0$ such that

$$P\left(|f_n(Z_n)| > C \frac{(a \log n)^b}{n^k}\right) = o(n^{-r}). \quad (6)$$

4.3 COROLLARY. *The order of approximation in (3) is $O_Z(n^{-(r+2)/2})$.*

Proof: According to Theorem 4.1 we have

$$\sup_{z_n} |f_n(z_n) - g_{n,r+1}(z_n)| = O(n^{-(r+2)/2}).$$

Let $L_{na} := \{z : |z| \leq \sqrt{a \log n}\}$. From Lemma 6.1 it follows that for suitable a

$$\sup_{z \notin L_{na}} |f_n(z)| = O(n^{-(r+2)/2}).$$

Similarly, we obtain from the obvious properties of the function $\phi(z)$ that

$$\sup_{z \notin L_{na}} |g_{n,r+1}(z)| = O(n^{-(r+2)/2}).$$

Now the assertion follows from

$$\frac{1}{\sigma_n} \phi(Z_n) \frac{1}{n^{(r+1)/2}} P_{n,r+1}(Z_n) = O_Z(n^{-(r+2)/2}).$$

□

5 Proofs

This section contains the proofs of the main results. We refer to notations and facts presented in section 4. All assertions are based on the assumption that the sequence (p_i) is contained in a closed subinterval of $(0, 1)$.

Let i and $j, i \neq j$, be fixed. Let us agree that symbols with index $n-1$ refer to summation over $k \neq i, 1 \leq k \leq n$, and that symbols with index $n-2$ refer to summation over $k \neq i, k \neq j, 1 \leq k \leq n$. This means e.g.

$$S_{n-1} := \sum_{j=1, j \neq i}^n X_j, \quad \sigma_{n-1}^2 := \sum_{j=1, j \neq i}^n p_j(1-p_j).$$

Our starting point are the obvious identities

$$E(X_i | S_n = s) = p_i \frac{P(S_{n-1} = s-1)}{P(S_n = s)} = p_i \frac{f_{n-1}(z_{n-1, s-1})}{f_n(z_{ns})}, \quad (7)$$

$$E(X_i X_j | S_n = s) = p_i p_j \frac{P(S_{n-2} = s-2)}{P(S_n = s)} = p_i p_j \frac{f_{n-2}(z_{n-2, s-2})}{f_n(z_{ns})}. \quad (8)$$

It is clear that the likelihood ratios f_{n-1}/f_n and f_{n-2}/f_n are uniformly bounded. We want to replace the likelihood ratios f_{n-1}/f_n and f_{n-2}/f_n by the ratios $g_{n-1,r}/g_{n,r}$ and $g_{n-2,r}/g_{n,r}$ of the corresponding asymptotic expansions. As a first step we are going to consider the latter ratios.

The following abbreviations will turn out to be useful:

$$\begin{aligned} c_{ni}^2 &= \frac{v_i}{\bar{\sigma}_n^2} & d_{ni} &= \frac{1-p_i}{\bar{\sigma}_n}, \\ c_{nij}^2 &= \frac{v_i + v_j}{\bar{\sigma}_n^2} & d_{nij} &= \frac{2-p_i-p_j}{\bar{\sigma}_n}. \end{aligned} \quad (9)$$

5.1 LEMMA. For $k = 1, 2$ and $r \geq 3$ the expansion

$$\frac{g_{n-k,r}(Z_{n-k})}{g_{nr}(Z_n)} = 1 + \frac{1}{n^{1/2}} A_k + \frac{1}{n} (B_k + V_k) + \frac{1}{n^{3/2}} (C_k + A_k V_k + W_k) + O_Z(n^{-2})$$

holds with

$$\begin{aligned} A_1 &= d_{ni} Z_n, & A_2 &= d_{nij} Z_n \\ B_1 &= \frac{d_{ni}^2 - c_{ni}^2}{2} (Z_n^2 - 1), & B_2 &= \frac{d_{nij}^2 - c_{nij}^2}{2} (Z_n^2 - 1) \\ C_1 &= \left(-\frac{d_{ni} c_{ni}^2}{2} + \frac{d_{ni}^3}{6} \right) (Z_n^3 - 3Z_n), & C_2 &= \left(-\frac{d_{nij} c_{nij}^2}{2} + \frac{d_{nij}^3}{6} \right) (Z_n^3 - 3Z_n) \end{aligned}$$

$$V_1 = -d_{ni}P'_{n1}$$

$$W_1 = d_{ni}(P'_{n1}P_{n1} - P'_{n2}) + \frac{c_{ni}^2}{2}Z_nP'_{n1} + \frac{d_{ni}^2}{2}P''_{n1} + \frac{3c_{ni}^2}{2}P_{n1} - \frac{\chi_{3i}}{\chi_{n3}}P_{n1}$$

and

$$V_2 = -d_{nij}P'_{n1}$$

$$W_2 = d_{nij}(P'_{n1}P_{n1} - P'_{n2}) + \frac{c_{nij}^2}{2}Z_nP'_{n1} + \frac{d_{nij}^2}{2}P''_{n1} + \frac{3c_{nij}^2}{2}P_{n1} - \frac{\chi_{3i} + \chi_{3j}}{\chi_{n3}}P_{n1}.$$

(By abuse of notation we simply denote $P_{nj} := P_{nj}(Z_n)$.)

Proof: Recall that

$$\frac{g_{n-k,3}(Z_{n-k,s-k})}{g_{n3}(Z_{ns})} = \frac{\sigma_n}{\sigma_{n-k}} \frac{\phi(Z_{n-k,s-k})}{\phi(Z_{ns})} \cdot \frac{1 + \sum_{j=1}^r \frac{1}{(n-k)^{j/2}} P_{n-k,j}(Z_{n-k,s-k})}{1 + \sum_{j=1}^r \frac{1}{n^{j/2}} P_{nj}(Z_{ns})}. \quad (10)$$

Let us start with expanding the first factor of (10)

$$\frac{\sigma_n}{\sigma_{n-k}} \frac{\phi(Z_{n-k})}{\phi(Z_n)}.$$

Using

$$\phi(x+h) = \phi(x) \left(1 + \sum_{k=1}^3 (-1)^k H_k(x) \frac{h^k}{k!} \right) + O(h^4) \quad (11)$$

we obtain by Lemma 6.5 and Lemma 6.6 and simple computations

$$\frac{\sigma_n}{\sigma_{n-k}} \frac{\phi(Z_{n-k})}{\phi(Z_n)} = 1 + \frac{A_k}{n^{1/2}} + \frac{B_k}{n} + \frac{C_k}{n^{3/2}} + O_Z(n^{-2}). \quad (12)$$

Now we turn to the second factor of (10)

$$\frac{1 + \sum_{j=1}^r \frac{1}{(n-k)^{j/2}} P_{n-k,j}(Z_{n-k})}{1 + \sum_{j=1}^r \frac{1}{n^{j/2}} P_{nj}(Z_n)}.$$

By Lemma 6.3 we obtain for the numerator

$$1 + \frac{1}{(n-k)^{1/2}} P_{n-k,1} + \frac{1}{n-k} P_{n-k,2} + \frac{1}{(n-k)^{3/2}} P_{n-k,3} + O_Z(n^{-2})$$

$$= 1 + \frac{1}{n^{1/2}} P_{n-k,1} + \frac{1}{n} P_{n-k,2} + \frac{1}{n^{3/2}} \left(P_{n-k,3} + \frac{k}{2} P_{n-k,1} \right) + O_Z(n^{-2}). \quad (13)$$

The denominator is expanded as $(1+x)^{-1} = 1 - x + x^2 - x^3 + O(x^4)$ which gives

$$\begin{aligned}
& \left(1 + \frac{1}{n^{1/2}}P_{n,1} + \frac{1}{n}P_{n,2} + \frac{1}{n^{3/2}}P_{n,3} + O_Z(n^{-2})\right)^{-1} \\
&= 1 - \frac{1}{n^{1/2}}P_{n,1} - \frac{1}{n}P_{n,2} - \frac{1}{n^{3/2}}P_{n,3} \\
&\quad + \frac{1}{n}P_{n,1}^2 + \frac{2}{n^{3/2}}P_{n,1}P_{n,2} - \frac{1}{n^{3/2}}P_{n,1}^3 + O_Z(n^{-2}) \\
&= 1 - \frac{1}{n^{1/2}}P_{n,1} + \frac{1}{n}(P_{n,1}^2 - P_{n,2}) + \frac{1}{n^{3/2}}(2P_{n,1}P_{n,2} - P_{n,3} - P_{n,1}^3) + O_Z(n^{-2}). \quad (14)
\end{aligned}$$

Next we multiply (13) and (14) and observe that by Lemma 6.9

$$P_{n-k,j}(Z_{n-k}) - P_{n,j}(Z_n) = O_Z(n^{-1/2}). \quad (15)$$

This results in the asymptotic expansion

$$\begin{aligned}
& \frac{1 + \sum_{j=1}^r \frac{1}{(n-k)^{j/2}}P_{n-k,j}(Z_{n-k,s-k})}{1 + \sum_{j=1}^r \frac{1}{n^{j/2}}P_{n,j}(Z_{ns})} \\
&= 1 + \frac{1}{n}\sqrt{n}(P_{n-k,1} - P_{n,1}) \\
&\quad + \frac{1}{n^{3/2}}\left(\frac{k}{2}P_{n,1} + \sqrt{n}(P_{n-k,2} - P_{n,2}) - \sqrt{n}(P_{n-k,1} - P_{n,1})P_{n,1}\right) + O_Z(n^{-2}) \\
&= 1 + \frac{V_k}{n} + \frac{W_k}{n^{3/2}} + O_Z(n^{-2}). \quad (16)
\end{aligned}$$

It is remarkable that the term of order $n^{-1/2}$ vanishes. The explicit form of the terms V_k and W_k follows from Lemma 6.9 by simple computations. \square

The expansions of Lemma 5.1 are polynomials in Z_n with uniformly bounded coefficients. This implies the following corollary. Recall that $L_{na} := \{z : |z| \leq \sqrt{a \log n}\}$.

5.2 COROLLARY. *The ratios $g_{n-1,r}/g_{n,r}$ and $g_{n-2,r}/g_{n,r}$, $r \geq 1$, are $1 + O_Z(n^{-1/2})$. They have uniformly bounded moments of all orders and they are uniformly bounded on L_{na} .*

Let us turn to the distances

$$\Delta_{nk,r}(z_n) := \frac{f_{n-k}(z_{n-k})}{f_n(z_n)} - \frac{g_{n-k,r}(z_{n-k})}{g_{nr}(z_n)}. \quad (17)$$

It is clear that the assertions of Corollary 5.2 are also valid for these distances.

5.3 LEMMA. *Let $K \subseteq \mathbb{R}$ be bounded and $r \geq 3$. Then*

$$(1) \sup_{z_n \in K} |\Delta_{nk,r}(z_n)| = O(n^{-(r+1)/2}).$$

(2) $E(|\Delta_{nk,r}(Z_n)|^m) = O(n^{-r/2})$ for any $m \in \mathbb{N}$.

Proof: From Lemma 6.2 it follows that

$$\Delta_{nk,r} \leq \frac{1}{f_n} \left(|f_n - g_{nr}| + \frac{g_{n-k,r}}{g_{nr}} |f_{n-k} - g_{n-k,r}| \right). \quad (18)$$

The ratios $g_{n-k,r}/g_{nr}$ are uniformly bounded on L_{na} and therefore on K . Since $\sqrt{n}f_n(z_n)$ for $z_n \in K$ is uniformly bounded away from zero assertion (1) follows from Corollary 4.3 .

For the proof of assertion (2) we split the expectation into two parts

$$E(|\Delta_{nk,r}(Z_n)|^m) = \int_{L_{na}} |\Delta_{nk,r}(Z_n)|^m dP + \int_{L'_{na}} |\Delta_{nk,r}(Z_n)|^m dP.$$

For the first part assertion (2) follows from (18) with the aid of Corollary 5.2 and Theorem 4.1 ,(2). For the second part we apply Lemma 6.1 and Corollary 5.2 . \square

The remainders in 5.1 fulfil (5) and (6). Together with Lemma 5.3 we obtain expansions of the likelihood ratios $f_{n-k,r}/f_{nr}$.

5.4 THEOREM. For $k = 1, 2$ the expansions

$$\frac{f_{n-k}(Z_{n-k})}{f_n(Z_n)} = 1 + \frac{1}{n^{1/2}} A_k + \frac{1}{n} (B_k + V_k) + \frac{1}{n^{3/2}} (C_k + A_k V_k + W_k) + r_n(Z_n).$$

hold with

$$E(|r_n(Z_n)|) = O(n^{-2}) \quad \text{and} \quad |r_n| = O_c(n^{-2}).$$

Now we are in the position to prove Theorem 2.4 .

Proof: (of Theorem 2.4) We keep the notation from sections 4 and 5.

The proof of the first assertion starts with

$$\begin{aligned} E\left(E(X_i^2|S_n) - E(X_i|S_n)^2\right) &= V(X_i) - E\left(\left(E(X_i|S_n) - E(X_i)\right)^2\right), \\ &= v_i - p_i^2 E\left(\left(\frac{f_{n-1}(Z_{n-1})}{f_n(Z_n)} - 1\right)^2\right). \end{aligned}$$

Apply Theorem 5.4 for $k = 1$. Denoting $X_1 := B_1 + V_1$ and $Y_1 := C_1 + A_1 V_1 + W_1$ we get

$$\left(\frac{f_{n-1}(Z_{n-1})}{f_n(Z_n)} - 1\right)^2 = \frac{1}{n} A_1 + \frac{1}{n^{3/2}} 2A_1 X_1 + \frac{1}{n^2} (X_1^2 + 2A_1 Y_1)^2 + r_n(Z_n)$$

where $E(|r_n(Z_n)|) = O(n^{-5/2})$. All terms are polynomials in Z_n . Taking expectations we note that $E(Z_n) = 0$, $E(Z_n^2) = 1$,

$$E(Z_n^3) = \text{cum}_3(Z_n) = \frac{\text{cum}_3(S_n - n\bar{p})}{\sigma_n^3} = \frac{n\bar{\chi}_{n3}}{\sigma_n^3} = \frac{1}{n^{1/2}} \frac{\bar{\chi}_{n3}}{\bar{\sigma}_n^3}$$

and

$$E(Z_n^4) = 3 + O(n^{-1/2}).$$

Then the assertion follows by simple calculations.

The proof of the second assertion starts with

$$\begin{aligned} & E\left(E(X_i X_j | S_n) - E(X_i | S_n)E(X_j | S_n)\right) \\ &= \text{Cov}(X_i, X_j) - E\left((E(X_i | S_n) - E(X_i))(E(X_j | S_n) - E(X_j))\right) \\ &= -p_i p_j E\left(\left(\frac{f_{n-1}^{(i)}(Z_{n-1}^{(i)})}{f_n(Z_n)} - 1\right)\left(\frac{f_{n-1}^{(j)}(Z_{n-1}^{(j)})}{f_n(Z_n)} - 1\right)\right). \end{aligned}$$

Apply again Theorem 5.4 for $k = 1$, but for different indices i and j separately. Using corresponding abbreviations as before we get

$$\begin{aligned} & \left(\frac{f_{n-1}^{(i)}(Z_{n-1}^{(i)})}{f_n(Z_n)} - 1\right)\left(\frac{f_{n-1}^{(j)}(Z_{n-1}^{(j)})}{f_n(Z_n)} - 1\right) \\ &= \frac{1}{n} A_1^{(i)} A_1^{(j)} + \frac{1}{n^{3/2}} (A_1^{(i)} X_1^{(j)} + A_1^{(j)} X_1^{(i)}) \\ &+ \frac{1}{n^2} (X_1^{(i)} X_1^{(j)} + A_1^{(i)} Y_1^{(j)} + A_1^{(j)} Y_1^{(i)})^2 + r_n(Z_n) \end{aligned}$$

where $E(|r_n(Z_n)|) = O(n^{-5/2})$. Again all terms are polynomials in Z_n . Taking expectations and remembering the moments of Z_n the assertion follows by simple calculations. \square

6 Auxiliary lemmas

We start with two easy probabilistic facts.

From Lemma 14.1. in Bhattacharya-Rao [1] it follows that

$$E(|Z_n|^m) \leq C_m < \infty, \quad n, m \in \mathbb{N}. \quad (19)$$

Moreover, we obtain the following bound for the tails of (Z_n) .

6.1 LEMMA. Let $L_{na} := \{z : |z| \leq \sqrt{a \log n}\}$. For any $r \geq 0$ there is a $a := a(r)$ such that

$$E(1_{Z_n \notin L_{na}} |Z_n|^m) = o(n^{-(r+1)/2}), \quad m \geq 0.$$

Proof: From Corollary 17.13 in Bhattacharya-Rao [1] we obtain the existence of a constant a such that

$$P(Z_n \notin L_{na}) = o(n^{-(r+1)}).$$

Applying the Cauchy-Schwarz inequality and (19) proves the assertion. \square

Now we turn to expansions. The first two lemmas are trivial.

6.2 LEMMA. If $a, b, c, d \in \mathbb{R}$ then

$$\left| \frac{a}{b} - \frac{c}{d} \right| \leq \frac{1}{b} \left(|a - c| + \frac{c}{d} |b - d| \right).$$

6.3 LEMMA.

$$\frac{1}{\sqrt{n-k}} = \frac{1}{n^{1/2}} + \frac{k}{2n^{3/2}} + O(n^{-2})$$

$$\frac{1}{n-k} = \frac{1}{n} + O(n^{-2})$$

Recall our notational convention. Let i and j , $i \neq j$, be fixed. All symbols with index $n-1$ refer to summation over $k \neq i$, $1 \leq k \leq n$, and all symbols with index $n-2$ refer to summation over $k \neq i, k \neq j$, $1 \leq k \leq n$.

6.4 LEMMA. Let (a_j) be a sequence of real numbers contained in a closed subinterval of $(0, 1)$. Then

$$\frac{\bar{a}_{n-1}}{\bar{a}_n} = 1 + \frac{1}{n} - \frac{a_i}{n\bar{a}_n} + O(n^{-2}),$$

$$\frac{\bar{a}_{n-2}}{\bar{a}_n} = 1 + \frac{2}{n} - \frac{a_i + a_j}{n\bar{a}_n} + O(n^{-2}).$$

Proof: The first assertion follows from

$$\begin{aligned}\frac{\bar{a}_{n-1}}{\bar{a}_n} &= \left(1 - \frac{1}{n}\right)^{-1} \left(1 - \frac{a_i}{n\bar{a}_n}\right) \\ &= \left(1 + \frac{1}{n} + O(n^{-2})\right) \left(1 - \frac{a_i}{n\bar{a}_n}\right) \\ &= 1 + \frac{1}{n} - \frac{a_i}{n\bar{a}_n} + O(n^{-2}).\end{aligned}$$

The second assertion follows from

$$\begin{aligned}\frac{\bar{a}_{n-1}}{\bar{a}_n} &= \left(1 - \frac{2}{n}\right)^{-1} \left(1 - \frac{a_i + a_j}{n\bar{a}_n}\right) \\ &= \left(1 + \frac{2}{n} + O(n^{-2})\right) \left(1 - \frac{a_i + a_j}{n\bar{a}_n}\right) \\ &= 1 + \frac{2}{n} - \frac{a_i + a_j}{n\bar{a}_n} + O(n^{-2}).\end{aligned}$$

□

An immediate consequence is the next lemma.

6.5 LEMMA.

$$\begin{aligned}\frac{\sigma_n^k}{\sigma_{n-1}^k} &= 1 + \frac{k c_{ni}^2}{2n} + O(n^{-2}), \\ \frac{\sigma_n^k}{\sigma_{n-2}^k} &= 1 + \frac{k c_{nij}^2}{2n} + O(n^{-2}).\end{aligned}$$

Let us turn to stochastic expansions.

6.6 LEMMA.

$$\begin{aligned}Z_{n-1} &= Z_n - \frac{1}{n^{1/2}} d_{ni} + \frac{1}{n} \frac{c_{ni}^2}{2} Z_n - \frac{1}{n^{3/2}} \frac{d_{ni} c_{ni}^2}{2} + O_Z(n^{-2}), \\ Z_{n-2} &= Z_n - \frac{1}{n^{1/2}} d_{nij} + \frac{1}{n} \frac{c_{nij}^2}{2} Z_n - \frac{1}{n^{3/2}} \frac{d_{nij} c_{nij}^2}{2} + O_Z(n^{-2}).\end{aligned}$$

Proof: The first assertion follows from Lemma 6.5 and

$$\begin{aligned}Z_{n-1} &= \frac{S_n - \mu_n}{\sigma_{n-1}} - \frac{1 - p_i}{\sigma_{n-1}} = \left(Z_n - \frac{d_{ni}}{\sqrt{n}}\right) \left(1 + \frac{c_{ni}^2}{2n}\right) + O_Z(n^{-2}) \\ &= Z_n - \frac{1}{n^{1/2}} d_{ni} + \frac{1}{n} \frac{c_{ni}^2}{2} Z_n - \frac{1}{n^{3/2}} \frac{d_{ni} c_{ni}^2}{2} + O_Z(n^{-2}).\end{aligned}$$

The second assertion is proved similarly. \square

6.7 LEMMA. *Hermite polynomials satisfy*

$$H_k(Z_{n-1}) = H_k(Z_n) - \frac{1}{n^{1/2}} d_{ni} H'_k(Z_n) + \frac{1}{n} \left(\frac{c_{ni}^2}{2} Z_n H'_k(Z_n) + \frac{d_{ni}^2}{2} H''_k(Z_n) \right) + O_Z(n^{-3/2}),$$

$$H_k(Z_{n-2}) = H_k(Z_n) - \frac{1}{n^{1/2}} d_{nij} H'_k(Z_n) + \frac{1}{n} \left(\frac{c_{nij}^2}{2} Z_n H'_k(Z_n) + \frac{d_{nij}^2}{2} H''_k(Z_n) \right) + O_Z(n^{-3/2}).$$

Proof: Apply Lemma 6.6 . \square

6.8 LEMMA. *We have*

$$w_{n-1,3} = w_{n3} \left(1 + \frac{\Delta_i w_{n3}}{n} \right) + O(n^{-2}) \text{ where } \Delta_i w_{n3} = -\frac{1}{2} + \frac{3}{2} c_{ni}^2 - \frac{\chi_{ki}}{\bar{\chi}_{nk}}$$

and

$$w_{n-2,3} = w_{n3} \left(1 + \frac{\Delta_{ij} w_{n3}}{n} \right) + O(n^{-2}) \text{ where } \Delta_{ij} w_{n3} = -1 + \frac{3}{2} c_{nij}^2 - \frac{\chi_{ki} + \chi_{kj}}{\bar{\chi}_{nk}}.$$

Proof: The first assertion follows from Lemma 6.4 by

$$\begin{aligned} w_{n-1,3} &= w_{n3} \frac{w_{n-1,3}}{w_{n3}} = w_{n3} \left(\frac{\bar{\sigma}_n^2}{\bar{\sigma}_{n-1}^2} \right)^{3/2} \frac{\bar{\chi}_{n-1,3}}{\bar{\chi}_{n3}} \\ &= w_{n3} \left(1 - \frac{1}{n} + \frac{v_i}{n \bar{\sigma}_n^2} \right)^{3/2} \left(1 + \frac{1}{n} - \frac{\chi_{3i}}{n \bar{\chi}_{n3}} \right) + O(n^{-2}) \\ &= w_{n3} + \frac{1}{n} w_{n3} \left(-\frac{3}{2} + \frac{3v_i}{2\bar{\sigma}_n^2} + 1 - \frac{\chi_{3i}}{\bar{\chi}_{n3}} \right) + O(n^{-2}). \end{aligned}$$

The second assertion is proved similarly. \square

6.9 LEMMA. *Edgeworth polynomials satisfy*

$$P_{n-1,k}(Z_{n-1}) - P_{nk}(Z_n) = -\frac{1}{n^{1/2}} d_{ni} P'_{nk}(Z_n) + O_Z(n^{-1}),$$

$$P_{n-2,k}(Z_{n-2}) - P_{nk}(Z_n) = -\frac{1}{n^{1/2}} d_{nij} P'_{nk}(Z_n) + O_Z(n^{-1}).$$

In particular

$$P_{n-1,1}(Z_{n-1}) - P_{n1}(Z_n) = -\frac{1}{n^{1/2}}d_{ni}P'_{n1}(Z_n) + \frac{1}{n}\left(\frac{c_{ni}^2}{2}Z_nP'_{n1}(Z_n) + \frac{d_{ni}^2}{2}P''_{n1}(Z_n) + \Delta_i w_{n3}P_{n1}(Z_n)\right) + O_Z(n^{-3/2}),$$

$$P_{n-2,1}(Z_{n-2}) - P_{n1}(Z_n) = -\frac{1}{n^{1/2}}d_{nij}P'_{n1}(Z_n) + \frac{1}{n}\left(\frac{c_{nij}^2}{2}Z_nP'_{n1}(Z_n) + \frac{d_{nij}^2}{2}P''_{n1}(Z_n) + \Delta_{ij}w_{n3}P_{n1}(Z_n)\right) + O_Z(n^{-3/2}).$$

Proof: Edgeworth polynomials consist of cumulants and Hermite polynomials. Hence the first assertion follows from Lemma 6.8 and Lemma 6.7. The second assertion follows by substituting the expressions from Lemma 6.8 and Lemma 6.7 into

$$P_{n-1,1}(Z_{n-1}) - P_{n1}(Z_n) = \frac{w_{n3}}{3!}\left(H_3(Z_{n-1}) - H_3(Z_n)\right) + \frac{w_{n-1,3} - w_{n3}}{3!}H_3(Z_{n-1}) + O_Z(n^{-3/2}).$$

□

References

- [1] R. N. Bhattacharya and R. Ranga Rao. *Normal approximation and asymptotic expansions*. Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley&Sons, Inc. XIV, 274 p., 1976.
- [2] P. Billingsley. *Convergence of probability measures*. 2nd ed. Wiley Series in Probability and Statistics. Chichester: Wiley. , 1999.
- [3] K. Hoffmann. *Approximation von stochastischen Integralgleichungen durch Martingaldifferenzschemata*. PhD thesis, University of Bayreuth, 1987.
- [4] H. Strasser. The covariance of cml-estimates in the rasch model when the number of item parameters is large. Technical report, Institute of Statistics and Mathematics, WU, 2011.
- [5] H. Strasser. Managing Edgeworth expansions with maxima. Technical report, Institute of Statistics and Mathematics, WU, 2011.
- [6] H. Strasser. Numerical studies for the rasch model with many item parameters. Technical report, Institute of Statistics and Mathematics, WU, 2011.