Version 2000-11-17 Ref: Proceedings. University of West Bohemia, vol. 4 (2000), Section 1: Natural and engineering sciences, University of West Bohemia, Pilsen, 2001, pp. 39–42.
Michal Friesl

AN ASYMPTOTIC EXPANSION FOR EXPECTATIONS OF FUNCTIONS OF SUMS VIA CONDITIONAL MOMENTS

MICHAL FRIESL

Abstract. For functions of sums of conditionally independent random variables an asymptotic expansion of their expectations is presented. As an example, the expansion of a Bayes risk is given.

Key words and phrases. Asymptotic expansion, Bayes risk.

1. Introduction

Let random variables X_n and functions g_n, $n \in \mathbf{N}$, be given, the variables X_n being asymptotically close to some value μ. Hurt (1986) presents a set of propositions on asymptotic expansions of $\mathrm{E}\, g_n(X_n)$ as $n \to \infty$ using the first q derivatives of g_n with remainders of the order $O(\mathrm{E}(X_n - \mu)^{q+1})$. If X_n is an arithmetic mean of n i.i.d. random variables and $\mu = \mathrm{E}\, X_1$, then the remainder is $O(1/n^{(q+1)/2})$.

In the present paper we assume that distributions of X_n's are defined conditionally given values of a random variable θ and that the asymptotic value μ depends on θ too. In a special case, X_n's will be cumulated sums of conditionally (given θ) i.i.d. random variables, and $\mu = \mu_\theta = \mathrm{E}[X_1 \mid \theta]$ will be their conditional expected value.

If we use an expansion from Hurt (1986) for each particular value of θ then the remainders depend on θ in an unknown way. To obtain expansions of unconditional moments by taking expectations of the conditional ones with respect to θ, we need a more detailed form of the remainder.

2. Propositions

First we adopt a form of the remainder given in Hurt (1986), Theorem 2.

Proposition 2.1. *Let $X_1, X_2, \dots$ be random variables with values in an interval $A \subset \mathbf{R}$ and $\mu \in A$. Let $g_n\colon A \to \mathbf{R}$, $n \in \mathbf{N}$, be functions having continuous derivatives of the order $(q+1)$ on $M = (\mu-\varepsilon, \mu+\varepsilon) \cap A$. Let $k \geq 0$ be an integer and let $\mathrm{E}\,|X_n - \mu|^{q+1+2k}$ exist for $n \in \mathbf{N}$. If for some $0 \leq u \leq k$, $a, a_0, a_1, \dots, a_{q+1} \geq 0$, and for all $n \in \mathbf{N}$*

$$
\begin{gathered}
|g_n^{(j)}(\mu)| \leq a_j, \quad j = 0, \dots, q, \\
|g_n^{(q+1)}(x)| \leq a_{q+1} \quad \text{on } M, \qquad |g_n(x)| \leq a|x|^u n^k \quad \text{on } M^c = A \setminus M,
\end{gathered}
\tag{2.1}
$$

then

$$
\begin{aligned}
\Big|\mathrm{E}\, g_n(X_n) - \sum_{j=0}^{q} \frac{g_n^{(j)}(\mu)}{j!}\, \mathrm{E}(X_n - \mu)^j\Big| \leq C\bigg(&\Big(\sum_{j=0}^{q+1} a_j \varepsilon^{j-q-1}\Big)\, \mathrm{E}\,|X_n - \mu|^{q+1} \\
&+ a\Big(1 + \frac{1 + \mu^{q+1+2k}}{\varepsilon^{q+1+2k}}\Big) n^k\, \mathrm{E}\,|X_n - \mu|^{q+1+2k}\bigg),
\end{aligned}
\tag{2.2}
$$

where C depends on q and k only.

Proof. Denoting F_n the distribution of X_n, we have

$$\mathrm{E}\, g_n(X_n) - \sum_{j=0}^{q} \frac{g_n^{(j)}(\mu)}{j!}\, \mathrm{E}(X_n-\mu)^j = I + J - \sum_{j=0}^{q} \frac{J_j}{j!} \tag{2.3}$$

with

$$I = \int_M g_n(x) - \sum_{j=0}^{q} \frac{g_n^{(j)}(\mu)}{j!}(x-\mu)^j \,\mathrm{d}F_n(x), \qquad J = \int_{M^c} g_n(x)\,\mathrm{d}F_n(x),$$

$$J_j = \int_{M^c} g_n^{(j)}(\mu)(x-\mu)^j \,\mathrm{d}F_n(x), \quad j = 0,\dots,q.$$

We estimate all terms involved in (2.3).

For $x \in M$ we use Taylor expansion of g_n around μ to write (with a suitable ξ_x between x and μ, and hence in M)

$$|I| = \left|\int_M \frac{g_n^{(q+1)}(\xi_x)}{(q+1)!}(x-\mu)^{q+1}\,\mathrm{d}F_n(x)\right| \le \frac{a_{q+1}}{(q+1)!}\,\mathrm{E}\,|X_n-\mu|^{q+1}.$$

Next we denote $v = q+1+2k$, $M_1 = M^c \cap \{x; |x| \ge 1\}$, $M_2 = M^c \cap \{x; |x| < 1\}$, and estimate J. Using bounds for g_n, we have

$$\begin{aligned}
|J| &\le an^k \int_{M^c} |x|^u \,\mathrm{d}F_n(x) \le an^k\Big(\int_{M_1} |x|^v\,\mathrm{d}F_n(x) + \int_{M_2} 1\,\mathrm{d}F_n(x)\Big)\\
&\le an^k\Big(2^{v-1}\int_{M_1}|x-\mu|^v\,\mathrm{d}F_n(x) + 2^{v-1}\int_{M_1}|\mu|^v\,\mathrm{d}F_n(x) + \int_{M_2} 1\,\mathrm{d}F_n(x)\Big)\\
&\le an^k\Big(2^{v-1}\,\mathrm{E}\,|X_n-\mu|^v + (2^{v-1}|\mu|^u + 1)\,\mathrm{P}[X_n \in M^c]\Big)\\
&\le an^k\left(2^{v-1} + \frac{2^{v-1}|\mu|^v+1}{\varepsilon^v}\right)\mathrm{E}\,|X_n-\mu|^v,
\end{aligned}$$

the probability $\mathrm{P}[X_n \in M^c] = \mathrm{P}[|X_n-\mu| > \varepsilon]$ being estimated by the Chebyshev inequality.

Using the Hölder and the Chebyshev inequalities, we get for the terms J_j ($j = 0,\dots,q$) the estimates

$$\begin{aligned}
|J_j| &\le a_j \int_{M^c} |x-\mu|^j\,\mathrm{d}F_n(x) \qquad (2.4)\\
&\le a_j\Big(\int_{M^c} 1\,\mathrm{d}F_n(x)\Big)^{(q+1-j)/(q+1)}\Big(\int_{M^c}|x-\mu|^{q+1}\,\mathrm{d}F_n(x)\Big)^{j/(q+1)}\\
&\le a_j\Big(\frac{\mathrm{E}\,|X_n-\mu|^{q+1}}{\varepsilon^{q+1}}\Big)^{(q+1-j)/(q+1)}\Big(\mathrm{E}\,|X_n-\mu|^{q+1}\Big)^{j/(q+1)} = a_j\frac{\mathrm{E}\,|X_n-\mu|^{q+1}}{\varepsilon^{q+1-j}}.
\end{aligned}$$

□

Particularly we will consider the special case $\mu_\theta = \mathrm{E}[X_1 \,|\, \theta]$ and suppose that X_n is a sum of conditionally (given θ) independent and identically distributed random variables.

Proposition 2.2. *Let $X_n = \frac{1}{n}\sum_{i=1}^n Y_i$ where given θ the variables Y_i are i.i.d. For each value of θ let the functions g_n, $n \in \mathbf{N}$, satisfy the assumptions of Proposition 2.1 with $\mu = \mu_\theta = \mathrm{E}[Y_i \mid \theta]$ and with constants $a, a_0, a_1, \dots, a_{q+1}, \varepsilon$ in (2.1) possibly depending on θ. For even naturals m let P_m denote a set of all products*

$\prod_{j=1}^{s} \mathrm{E}[|Y_1 - \mu_\theta|^{i_j} \mid \theta]$ *where* $1 \le s \le m/2$ *and integers* $i_1, \ldots, i_s \ge 2$ *are such that* $\sum_{j=1}^{s} i_j = m$*, and for* m *odd let* $P_m = \{1 + p, p \in P_{m+1}\}$*. If there exist the expectations*

$$\mathrm{E}\Big(\Big(1 + \frac{1 + |\mu_\theta|^{q+1+2k}}{\varepsilon_\theta^{q+1+2k}}\Big) p_{1\theta}\Big) \qquad \text{for all } p_{1\theta} \in P_{q+1+2k} \tag{2.5}$$

and

$$\mathrm{E}(a_{j\theta}\varepsilon_\theta^{j-q-1} p_{2\theta}) \qquad \text{for all } p_{2\theta} \in P_{q+1}, \quad j = 0, \ldots, q+1, \tag{2.6}$$

then

$$\mathrm{E}\, g_n(X_n) = \mathrm{E} \sum_{j=0}^{q} \frac{g_n^{(j)}(\mu_\theta)}{j!} \mathrm{E}[(X_n - \mu_\theta)^j \mid \theta] + O(1/n^{(q+1)/2}) \qquad \text{as } n \to \infty. \tag{2.7}$$

Proof. We have

$$\Big|\mathrm{E}\, g_n(X_n) - \mathrm{E} \sum_{j=0}^{q} \frac{g_n^{(j)}(\mu_\theta)}{j!} \mathrm{E}[(X_n - \mu_\theta)^j \mid \theta]\Big| \le \mathrm{E}\Big|\mathrm{E}[g_n(X_n) \mid \theta] - \sum_{j=0}^{q} \frac{g_n^{(j)}(\mu_\theta)}{j!} \mathrm{E}[(X_n - \mu_\theta)^j \mid \theta]\Big|,$$

and we apply (2.2). The order of the convergence (2.7) follows from the fact that for an arithmetic mean $\overline{Z}$ of n zero mean i.i.d. random variables $Z_1, \ldots, Z_n$ and each even positive integer m

$$\mathrm{E}\,|\overline{Z}|^m = \mathrm{E}\Big(\sum_{i=1}^{n} Z_i/n\Big)^m \le K n^{-m/2} \sum \mathrm{E}\, Z_1^{i_1} \ldots \mathrm{E}\, Z_1^{i_{m/2}},$$

where the last sum is over all $(i_1, \ldots, i_{m/2})$ with components from $\{0, 2, 3, \ldots, m\}$ such that $\sum_{\ell=1}^{m/2} i_\ell = m$, and K is a constant depending on m only. For m odd,

$$\mathrm{E}\,|\overline{Z}|^m = \mathrm{E}\,|\overline{Z}|^m I_{[|\overline{Z}| < n^{-1/2}]} + \mathrm{E}\,|\overline{Z}|^m I_{[|\overline{Z}| \ge n^{-1/2}]} \le n^{-m/2} + n^{1/2}\,\mathrm{E}\,\overline{Z}^{m+1},$$

the second term in the sum being estimated in the same way as the integral in (2.4) with $x - \mu = z$, $j = q = m$, $\varepsilon = n^{-1/2}$. □

In some applications, $\varepsilon_\theta = \mu_\theta/2$ is suitable and the (conditional) moments of Y_1 are bounded by linear combinations of powers of its first moment. Then we can replace (2.5) and (2.6) as follows.

Proposition 2.3. *Let* $X_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ *where given* θ *the variables* Y_i *are i.i.d. For each value of* θ *let the functions* g_n*,* $n \in \mathbf{N}$*, satisfy the assumptions of Proposition 2.1 with* $\mu = \mu_\theta = \mathrm{E}[Y_i|\theta] > 0$*,* $\varepsilon = \varepsilon_\theta = \mu_\theta/2$*, and with constants* $a, a_0, a_1, \ldots, a_{q+1}$ *in (2.1) possibly depending on* θ*. Let moments of* Y_1 *satisfy* $\mathrm{E}[|Y_1^j| \mid \theta] \le \sum_{i=1}^{j} c_i \mu_\theta^i$ *with some constants* $c_1, \ldots, c_j > 0$*,* $j = 1, \ldots, q + \delta + 1 + 2k$ *(where* $\delta = 0$ *for* q *odd and* $\delta = 1$ *for* q *even). If the expectations* $\mathrm{E}\, a_\theta \mu_\theta^i$ *for* $i = -(q + \delta + 2k), \ldots, q + \delta + 1 + 2k$ *and* $\mathrm{E}\, a_{j\theta}\mu_\theta^i$ *for* $i = j - q - \delta, \ldots, j + \delta$ *(*$j = 0, \ldots, q+1$*) exist then (2.7) holds.*

3. Example

Let us express the Bayes risk for a Bayes estimator of the expected number e of trials preceeding the first success in a sequence of Bernoulli trials with a success probability p.

Observing I successes during n trials and assuming a prior beta distribution $\mathrm{B}(r,s)$ for p (i.e. the natural conjugate system distribution), the Bayes estimator (under the quadratic loss) of $e=(1-p)/p$ is

$$\widehat{e}=\mathrm{E}(e\mid I)=(n-I+s)/(I+r-1).$$

For the associated Bayes risk we have

$$\mathrm{BR}\,\widehat{e}=\mathrm{E}(e-\widehat{e})^2=\mathrm{E}\operatorname{var}(e\mid I)=\mathrm{E}\,\frac{(n-I+s)(n+r+s-1)}{(I+r-1)^2(I+r-2)}=\frac{1}{n}\,\mathrm{E}\,g_n(I/n)$$

with

$$g_n(i)=(1-i+s/n)(1+(r+s-1)/n)(i+(r-1)/n)^{-2}(1+(r-2)/n)^{-1}. \tag{3.1}$$

An asymptotic expansion of $\mathrm{E}\,g_n(I/n)$ for $n\to\infty$ is obtained by conditioning on p and expanding around $\mathrm{E}(I/n\mid p)=p$. With regard to assumptions of Proposition 2.1 we have $q=0$, $k=3$, $g_n(i)\le const\cdot(r-2)^{-3}n^3$. We take $\mu=p$, $\varepsilon=p/2$, hence $g_n(p)$ and g_n' on $(p/2,3p/2)$ are bounded for all n by a multiple of p^{-3} and p^{-4}, respectively. Therefore, Proposition 2.3 can be applied. For $r>7$ the expectations $\mathrm{E}\,p^i$ for $i=-7,\dots,8$, $\mathrm{E}\,p^{-3}p^i$ for $i=-1,0$, and $\mathrm{E}\,p^{-4}p^i$ for $i=0,1,2$ are finite and we can write $\mathrm{E}\,g_n(I/n)=\mathrm{E}\,g_n(p)+O(1/\sqrt{n})$, which yields

$$\mathrm{BR}\,\widehat{e}=\frac{1}{n}\,\mathrm{E}\,\frac{1-p}{p^3}+O(1/n^{3/2})=\frac{1}{n}\,\frac{s(r+s-1)(r+s-2)}{(r-1)(r-2)(r-3)}+O(1/n^{3/2}).$$

In fact, under the same assumption $r>7$ this formula can be written with $O(1/n^2)$ instead of $O(1/n^{3/2})$ — if the expansion is made with $q=1$.

Resumé

Příspěvek uvádí asymptotický rozvoj středních hodnot náhodných veličin, které jsou funkcemi součtů podmíněně nezávislých náhodných veličin. Jako příklad je uveden rozvoj bayesovského rizika.

References

1. J. Hurt (1986), Asymptotic expansions for moments of functions of stochastic processes and their applications, *Statist. Decisions* **4**, no. 2–3, 251–271.

Mgr. Michal Friesl, Ph.D.

Department of Mathematics
Faculty of Applied Sciences
University of West Bohemia

Address: Univerzitní 22, 306 14 Pilsen, Czech Republic
E-mail: friesl@kma.zcu.cz
URL: http://home.zcu.cz/~friesl
http://www.kma.zcu.cz/Friesl