



HAL
open science

Asymptotic distribution of density ratios

Sabrina Antonelli, Giuliana Regoli

► **To cite this version:**

Sabrina Antonelli, Giuliana Regoli. Asymptotic distribution of density ratios. *Statistics and Probability Letters*, 2009, 79 (3), pp.289. 10.1016/j.spl.2008.08.009 . hal-00504133

HAL Id: hal-00504133

<https://hal.science/hal-00504133>

Submitted on 20 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Asymptotic distribution of density ratios

Sabrina Antonelli, Giuliana Regoli

PII: S0167-7152(08)00400-8
DOI: 10.1016/j.spl.2008.08.009
Reference: STAPRO 5191

To appear in: *Statistics and Probability Letters*

Received date: 14 September 2007

Revised date: 14 August 2008

Accepted date: 20 August 2008

Please cite this article as: Antonelli, S., Regoli, G., Asymptotic distribution of density ratios. *Statistics and Probability Letters* (2008), doi:10.1016/j.spl.2008.08.009

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Asymptotic distribution of density ratios

Antonelli Sabrina Regoli Giuliana*

Dipartimento di Matematica e Informatica

University of Perugia, Italy

August 14, 2008

Abstract

We find limit theorems regarding asymptotic distribution of density ratios. These results are applied to analyze the goodness of classical approximations. Furthermore these theorems are used in statistical inference procedures to find asymptotic distribution of likelihood ratios.

Keywords: Approximation of distribution, relative error, asymptotic local normality, likelihood ratio, Polya's distribution.

*Corresponding author: Giuliana Regoli, Dipartimento di Matematica e Informatica, Università di Perugia, Via Vanvitelli 1, 06100 Perugia, Italy. e-mail: regoli@dipmat.unipg.it, tel.: +39-075-5855022. Sabrina Antonelli, Ple Caduti dei Lager 13, 06024 Gubbio (PG), Italy. e-mail: sabrina_antonelli@yahoo.it, tel: +39-075-9277529.

1 Introduction

This paper focuses on asymptotic behavior of density ratios. Two kinds of random variables $\{X_n\}$ are considered: absolutely continuous random variables with density function $f_{X_n}(x)$ and \mathbb{Z} -valued variables with probability function $p_{X_n}(x)$ and density function defined as $f_{X_n}(x) = p_{X_n}(\lfloor x \rfloor)$. Let $\mu_n = E(X_n)$ and $\sigma_n^2 = V(X_n)$ be finite; the variables $Z_n = \frac{X_n - \mu_n}{\sigma_n}$ have density functions $f_{Z_n}(z) = f_{X_n}(\mu_n + z\sigma_n)\sigma_n$. We consider sequences which locally converge to normal distribution, that is:

$$\lim_{n \rightarrow \infty} f_{Z_n}(z) = \varphi(z), \quad (1)$$

where $\varphi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}$. In Section 2 we find some limit theorems, concerning asymptotic distribution and moment convergence of density ratio or probability ratio. In the following sections these results are applied to find asymptotic distribution of relative error and asymptotic distribution of likelihood ratios for not independent sampling. From the same results, we also derive explicit formulas for evaluating the goodness of the well known approximations: the Binomial-Hypergeometric approximation (see, for example, Sródka, 1963, and Ord, 1968), Poisson-Binomial, Binomial-Hypergeometric and Poisson-Hypergeometric approximations (e. g. Burr, 1973), Negative Binomial - Polya's distribution. Other approaches to similar approximations can be found in Johnson et al. (1992). Our work gives a unifying approach to study all these approximations and provides us with the asymptotic mean relative

error, as already done in Antonelli and Regoli (2005) for the Poisson-Binomial approximation. Furthermore in the Appendix we prove a theorem of local asymptotic normality for Polya's distribution.

2 Limit theorems

Consider two sequences $\{X_n\}$ and $\{Y_n\}$, with $E(X_n) = \mu_n$, $E(Y_n) = \tilde{\mu}_n$, $V(X_n) = \sigma_n^2$ and $V(Y_n) = \tilde{\sigma}_n^2$. Let $g_n(z) = \frac{f_{Y_n}(\mu_n + z\sigma_n)}{f_{X_n}(\mu_n + z\sigma_n)}$ and $g(z) = re^{\frac{1}{2}z^2(1-r^2)}$, then the following Lemma holds.

Lemma 1 *Suppose $\lim_{n \rightarrow \infty} \frac{\sigma_n}{\tilde{\sigma}_n} = r$, $\lim_{n \rightarrow \infty} \frac{\mu_n - \tilde{\mu}_n}{\tilde{\sigma}_n} = 0$. Moreover, for every sequence of real numbers z_n converging to z , suppose $\lim_{n \rightarrow \infty} f_{X_n}(\mu_n + z_n\sigma_n)\sigma_n = \varphi(z)$ and $\lim_{n \rightarrow \infty} f_{Y_n}(\tilde{\mu}_n + z_n\tilde{\sigma}_n)\tilde{\sigma}_n = \varphi(z)$. Then $\lim_{n \rightarrow \infty} g_n(z_n) = g(z)$. Therefore $g_n(z)$ uniformly converges to $g(z)$ on compact sets.*

Proof. $\lim_{n \rightarrow \infty} f_{Y_n}(\mu_n + z_n\sigma_n)\tilde{\sigma}_n = \lim_{n \rightarrow \infty} f_{Y_n}(\tilde{\mu}_n + (\frac{\mu_n - \tilde{\mu}_n + z_n\sigma_n}{\tilde{\sigma}_n})\tilde{\sigma}_n)\tilde{\sigma}_n = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2r^2}$, since $\lim_{n \rightarrow \infty} z_n \frac{\sigma_n}{\tilde{\sigma}_n} = zr$ and $\lim_{n \rightarrow \infty} \frac{\mu_n - \tilde{\mu}_n}{\tilde{\sigma}_n} = 0$. Hence we have: $\lim_{n \rightarrow \infty} \frac{f_{Y_n}(\mu_n + z_n\sigma_n)}{f_{X_n}(\mu_n + z_n\sigma_n)} = \lim_{n \rightarrow \infty} \frac{f_{Y_n}(\mu_n + z_n\sigma_n) \frac{\sigma_n \tilde{\sigma}_n}{\sigma_n \tilde{\sigma}_n}}{f_{X_n}(\mu_n + z_n\sigma_n) \sigma_n} = re^{\frac{1}{2}z^2(1-r^2)}$. \square

Furthermore we have the following theorem.

Theorem 2 *Assume the conditions of Lemma 1. Let $Z_n = \frac{X_n - \mu_n}{\sigma_n}$. Then the sequence $\{g_n(Z_n)\}$ converges in distribution to $\{g(Z)\}$, where Z is a standard normal random variable. Moreover $\lim_{n \rightarrow \infty} E(g_n(Z_n)) = E(g(Z))$.*

Proof. By Lemma 1, $g_n(z)$ uniformly converges to $g(z)$ on compact sets.

Therefore the convergence in distribution of the sequence $\{g_n(Z_n)\}$ to $g(Z)$

follows from the observation after Theorem 5.5, p. 34, in Billingsley (1968).

In fact, for the condition $\lim_{n \rightarrow \infty} f_{X_n}(\mu_n + z_n \sigma_n) \sigma_n = \varphi(z)$, the sequence of variables $\{Z_n\}$ converges in distribution to the variable Z ; moreover the function g is everywhere continuous and the sequence $\{g_n\}$ uniformly converges on compact sets. Now let F_n be the distribution function of Z_n and let F'_n be the distribution function of $Z'_n = \frac{Y_n - \mu_n}{\sigma_n}$, with $E(Z'_n) = m_n = \frac{\bar{\mu}_n - \mu_n}{\sigma_n}$ and $V(Z'_n) = \frac{\bar{\sigma}_n^2}{\sigma_n^2}$. The variables $g_n(Z_n)$ are uniformly integrable: in fact, since g_n are uniformly bounded on compact sets, it suffices to prove that

$\sup_n \int_{|z|>a} g_n(z) dF_n(z)$ is small for a sufficiently large. We have $\int_{|z|>a} g_n(z) dF_n(z) = \int_{|z|>a} dF'_n(z) = \text{Prob}(|Z'_n| > a) \leq \text{Prob}(|Z'_n - m_n| > a - |m_n|)$, and, by Chebyshev's inequality, we get $\int_{|y|>a} dF'_n(z) \leq \frac{\bar{\sigma}_n^2}{(a - |m_n|)^2 \sigma_n^2} \leq \frac{2C}{a^2}$, for a sufficiently large and C such that $|\frac{\bar{\sigma}_n^2}{\sigma_n^2}| < C$. Therefore the uniform integrability follows. Thus, by using, for example, Theorem 5.4, p. 32, in Billingsley (1968), it follows that

$\lim_{n \rightarrow \infty} E(g_n(Z_n)) = E(g(Z))$. \square

Remark The density function $f_{g(Z)}$, with $r > 0$ and $r \neq 1$, can be found by using the fact that Z^2 has chi-square distribution with one degree of freedom. Thus we have: $f_{g(Z)}(t) = \frac{r^{1/(1-r^2)}}{\sqrt{(1-r^2)\pi \ln(t/r)}} t^{-(2-r^2)/(1-r^2)}$. For $0 < r < 1$, the support is $t > r$, otherwise, for $r > 1$, the support is $0 < t < r$. Furthermore $E(g(Z)) = 1$, while the variance is finite only for $r > \frac{\sqrt{2}}{2}$ and it is

$V(g(Z)) = \frac{r^2}{\sqrt{(2r^2-1)}} - 1$.

In the previous theorem no relationships between X_n and Y_n are considered. By adding relationships other results can be found.

Put $V(Y_n|X_n) = \tilde{\sigma}_n(X_n)$ and $\tilde{g}_n(z) = \frac{f_{Y_n|X_n}(\mu_n + z\sigma_n|\mu_n + z\sigma_n)}{f_{X_n}(\mu_n + z\sigma_n)}$, then we have the following theorem.

Theorem 3 *Assume that $\frac{\sigma_n}{\tilde{\sigma}_n(X_n)}$ converges in probability to $r \neq 0$, $f_{X_n}(\mu_n + z_n\sigma_n)\sigma_n$ uniformly converges to a density function $f(z)$ on compact sets and $f_{Y_n|X_n}(X_n|X_n)\tilde{\sigma}_n(X_n)$ converges in probability to a real number $k \neq 0$. If $Z_n = \frac{X_n - \mu_n}{\sigma_n}$, then $\{\tilde{g}_n(Z_n)\}$ converges in distribution to $\frac{kr}{f(Z)}$, where Z has density function f .*

Proof. Set $w_n(z) = f_{X_n}(\mu_n + z\sigma_n)\sigma_n$. Since $\lim_{n \rightarrow \infty} w_n(z) = f(z)$, the sequences of variables $\{Z_n\}$ converges in distribution to Z , whose density function is f . Furthermore the convergence of $w_n(z)$ to f is uniform on compact sets, therefore $w_n(Z)$ converges in distribution to $f(Z)$. On the other hand, if $v_n(Z_n) = f_{Y_n|\mu_n + Z_n\sigma_n}(\mu_n + Z_n\sigma_n|\mu_n + Z_n\sigma_n)\tilde{\sigma}_n(\mu_n + Z_n\sigma_n)$, $v_n(Z_n)$ and $f_{Y_n|X_n}(X_n|X_n)\tilde{\sigma}_n(X_n)$ are identically distributed. Therefore $\tilde{g}_n(Z_n) = \frac{v_n(Z_n)\sigma_n}{w_n(Z_n)\tilde{\sigma}_n(\mu_n + Z_n\sigma_n)}$ converges in distribution to $\tilde{g}(Z) = \frac{kr}{f(Z)}$.

Remark If $f(z) = \varphi(z)$ and $k = \frac{1}{\sqrt{2\pi}}$, then $(2\ln(\tilde{g}(Z)) - 2\ln r)$ has chi-square distribution with one degree of freedom.

3 Approximation of distributions

The ratio $g_n(z)$ can be used for studying the relative error in the approximation of density or probability functions. When Y_n approximate X_n , the relative error at x is $|\frac{f_{Y_n}(x) - f_{X_n}(x)}{f_{X_n}(x)}| = |\frac{f_{Y_n}(x)}{f_{X_n}(x)} - 1|$. Suppose $\{X_n\}$ and $\{Y_n\}$

are two sequences which satisfy the conditions of Lemma 1. If x is written as $x(n, z) = \mu_n + z\sigma_n$, for a suitable real number z , the relative error can be written in terms of $g_n(z)$:

$$h_n(z) = \left| \frac{f_{Y_n}(\mu_n + z\sigma_n)}{f_{X_n}(\mu_n + z\sigma_n)} - 1 \right| = |g_n(z) - 1|.$$

Let $h(z) = |g(z) - 1|$, then $\lim_{n \rightarrow \infty} h_n(z) = h(z)$, for every z . Set $Z_n = \frac{X_n - \mu_n}{\sigma_n}$ and suppose Z is a standard normal random variable. Since $\{g_n(Z_n)\}$ is uniformly integrable and converges in distribution to $g(Z)$, then also $\{h_n(Z_n)\}$ is uniformly integrable and converges in distribution to $h(Z)$. Therefore, $\lim_{n \rightarrow \infty} E(h_n(Z_n)) = E(h(Z))$. In particular, the asymptotic mean relative error is:

$$E(h(Z)) = 4 |\Phi(y_0) - \Phi(y_0 r)|, \text{ with } y_0 = \sqrt{\frac{-2 \ln r}{1 - r^2}} \text{ and } r = \lim_{n \rightarrow \infty} \frac{\sigma_n}{\tilde{\sigma}_n}, \quad (2)$$

where Φ is the standard normal distribution function. Obviously an approximation is better when r is closer to one.

In Antonelli and Regoli (2005) this relative error was already studied for the Poisson-binomial approximation. In that case X_n is a binomial random variable with parameters n and p and Y_n is a Poisson variable with parameter $\lambda = np$. Given p , these sequences satisfy the conditions of Lemma 1, with $r = \sqrt{1 - p}$ (e. g. by Theorem 3 in Feller, 1950, vol. II, p. 490). In the same way we can study the goodness of the approximation of a distribution to another when both densities are asymptotically normal.

In Table 1, we report the distributions and the parameters used in well known

approximations.

Distribution	$p(x)$	parameters
Poisson	$\frac{\lambda^x}{x!}e^{-\lambda}$	λ
Binomial	$\binom{n}{x}p^x(1-p)^{n-x}$	n, p
Negative Binomial	$\binom{m-1+x}{m-1}q^m(1-q)^x$	m, q
Hypergeometric	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$	N, M, n
Polya's distribution	$\binom{n}{x} \frac{B(x+\frac{M}{k}, n-x+\frac{N-M}{k})}{B(\frac{M}{k}, \frac{N-M}{k})}$	N, M, n, k

Table 1: $\lambda > 0$, $0 < p < 1$, $0 < q < 1$, n, m, M, N , and k are positive

integers. B is the beta function.

In the examples outlined in the sequel formula (2) gives the asymptotic mean relative error, computing only r .

Binomial - Hypergeometric approximation. Given $\alpha, p \in (0; 1)$, for every natural number N , set $n = \lfloor \alpha N \rfloor$ and $M = \lfloor pN \rfloor$. Consider the sequence of hypergeometric variables $\{X_N\}$, with parameters N, M and n , and the sequence of binomial variables $\{Y_N\}$, with parameters n and p . The sequences $\{X_N\}$ and $\{Y_N\}$ satisfy the conditions of Lemma 1 as N tends to infinity, with $r = \sqrt{1 - \alpha}$.

Poisson - Hypergeometric approximation. Given $\alpha, p \in (0; 1)$, for every natural number N set $n = \lfloor \alpha N \rfloor$ and $M = \lfloor pN \rfloor$. Consider the sequence of hypergeometric variables $\{X_N\}$, with parameters N, M and n , and the sequence of Poisson variables $\{Y_N\}$ with parameter $\lambda = \alpha p N$. The two sequences

$\{X_N\}$ and $\{Y_N\}$ satisfy the conditions of Lemma 1 as N tends to infinity, with $r = \sqrt{(1 - \alpha)(1 - p)}$.

Poisson - Negative Binomial approximation. The Negative Binomial distribution with parameters m and q may be approximated by the Poisson distribution with parameter $\lambda = m \frac{1-q}{q}$ if q and m are large. Given q , as m tends to infinity, conditions of Lemma 1 are satisfied with $r = \frac{1}{\sqrt{q}}$.

Negative Binomial - Polya approximation. The Polya's distribution with parameters N , $M = \lfloor pN \rfloor$, k and n may be approximated by the Negative Binomial distribution with parameters $m = \lfloor \frac{pN}{k} \rfloor$ and $q = \frac{1}{1+\alpha k}$, if n and N are large and p is small (see, for example, Feller [1950], vol. I, p.131). Given $\alpha, p \in (0; 1)$ and $k = \{0, 1, 2, \dots\}$, for every natural number N set $n = \lfloor \alpha N \rfloor$ and $M = \lfloor pN \rfloor$. As N tends to infinity, the Polya's distribution tends to Normal distribution (the proof is provided in the Appendix). Polya's and Negative Binomial distributions satisfy the conditions of Lemma 1, with $r = \sqrt{1 - p}$.

The expression of r is given for each approximation in Table 2.

Approximation	r
Binomial - Hypergeometric	$\sqrt{1 - \alpha}$
Poisson - Hypergeometric	$\sqrt{(1 - \alpha)(1 - p)}$
Poisson - Negative Binomial	$\frac{1}{\sqrt{q}}$
Negative Binomial - Polya	$\sqrt{1 - p}$

Table 2: the expression of r for each approximation.

Formula (2) and expression of r suggest comparison among approximations. Given a hypergeometric variable X of parameters N , $M = \lfloor pN \rfloor$ and $n = \lfloor \alpha N \rfloor$ the binomial approximation with parameters n and p is better than Poisson approximation with parameter $\lambda = \alpha p N$, since $\sqrt{1 - \alpha} > \sqrt{(1 - \alpha)(1 - p)}$. In particular, for $\alpha < 0.1$, the mean relative error is less than 0.06 (see Table 5). Since Hypergeometric distribution is unchanged by interchanging M and n , the approximation improves by using a binomial with parameters M and α when $p < \alpha$. Furthermore when α and p are both close to 1, it is better to approximate $n - X$ by using a binomial distribution with parameters $\lfloor \min((1 - p)N; (1 - \alpha)N) \rfloor$ and $\max((1 - p); (1 - \alpha))$, as pointed out also in Johnson et al., pag. 258.

Numerical results concerning Poisson-binomial approximation are provided in Antonelli and Regoli (2005). Numerical results for the other approximations are shown in the following tables.

Binomial - Hypergeometric approximation

x	Binomial $n = 6$ and $p = 0.5$	Hypergeometric $N = 60, M = 30, n = 6$	Relative error	$h(z)$
0	0.0156	0.0119	0.3174	0.3167
1	0.0938	0.0854	0.0978	0.0975
2	0.2344	0.2381	0.0157	0.0161
3	0.3125	0.3293	0.0509	0.0513
4	0.2344	0.2381	0.0157	0.0161
5	0.0938	0.0854	0.0978	0.0975
6	0.0156	0.0119	0.3174	0.3167

Table 4: relative error and pointwise approximation $h(z)$, with $z = \frac{x-\mu}{\sigma}$,

$$\mu = 3 \text{ and } \sigma^2 = \frac{81}{59}.$$

α	$N = 50$	$N = 80$	$N = 100$	$E(h(Z))$
0.05	0.0204	0.0194	0.0256	0.0248
0.10	0.0526	0.0515	0.0518	0.0510
0.20	0.1095	0.1042	0.1049	0.1079

Table 5: mean relative error and asymptotic mean relative error $E(h(Z))$.

Binomial with $n = \alpha N$, $p = 0.5$ and Hypergeometric with N , $M = 0.5N$,

$$n = \alpha N.$$

Poisson - Hypergeometric approximation

α	$N = 50$	$N = 80$	$N = 100$	$E(h(Z))$
0.05	0.0566	0.0670	0.0722	0.0758
0.10	0.0967	0.1084	0.1154	0.1019
0.20	0.1781	0.1625	0.1525	0.1586

Table 6: mean relative error and asymptotic mean relative error $E(h(Z))$.

Poisson with $\lambda = 0.1\alpha N$ and Hypergeometric with N , $M = 0.1N$, $n = \alpha N$.

Poisson - Negative Binomial approximation

q	$m = 50$	$m = 80$	$m = 100$	$E(h(Z))$
0.80	0.1085	0.1083	0.1081	0.1079
0.90	0.0522	0.0516	0.0510	0.0510
0.95	0.0259	0.0254	0.0253	0.0248

Table 7: mean relative error and asymptotic mean relative error $E(h(Z))$.

Poisson with $\lambda = m \frac{1-q}{q}$ and Negative Binomial with m , q .

Negative Binomial - Polya approximation

p	$N = 50$	$N = 80$	$N = 100$	$E(h(Z))$
0.05	0.2358	0.0298	0.0280	0.0248
0.10	0.0578	0.0545	0.0525	0.0510
0.20	0.1155	0.1095	0.1098	0.1079

Table 8: mean relative error and asymptotic mean relative error $E(h(Z))$.

Negative Binomial with $m = \lfloor \frac{pN}{k} \rfloor$ and $q = \frac{1}{1+0.3}$ and Polya's distribution

with N , $M = pN$, $n = 0.3N$, $k = 1$.

4 Likelihood ratio

Theorem 3 may be used in statistical inference procedures to find the asymptotic distribution of a likelihood ratio. Suppose a sample of size n has a distribution which depends on an unknown parameter θ . Consider a sufficient statistic X_n with density (or probability function) $f_{X_n}(x; \theta)$. Let $\hat{\theta}_n$ be a function of X_n and let Y_n be a random variable whose conditional density (or probability function), given X_n , is $f_{X_n}(x; \hat{\theta}_n)$. We are tempted to use the last result of Section 2 in order to find the limit distribution of the likelihood ratio $\hat{g}_n(x) = \frac{f_{X_n}(x; \hat{\theta}_n)}{f_{X_n}(x; \theta)}$. For instance, it is straightforward to check that the assumptions of Theorem 3 are verified when X_n is the observation sum of independent bernoulli variables, with unknown parameter $\theta = p$, and $\hat{\theta}_n$ is the maximum likelihood estimate of p , that is the sample mean. In this case Theorem 3 provides us with another proof of the well known result that $2\ln(\hat{g})$ is asymptotically distributed as a chi-squared with one degree of freedom. But Theorems 3 may be used also when observations are not independent and identically distributed. Here, as a particular case, we shall analyze the sample observations coming from a Polya's urn scheme, with N, k, n known and unknown parameter $\theta = \frac{M}{N}$. The observations sum X_n is a sufficient statistic for θ .

Setting $\hat{\theta}_n = \frac{X_n}{n}$, Theorem 3 may be applied, as it is proved in the following proposition.

Proposition 4 *Let $\beta > 1$, set X_n a Polya's random variable with parameters $N = \beta n$, pN , k and n . Furthermore suppose Y_n , given X_n , is conditionally distributed as a Polya's random variable with parameters N , $\hat{\theta}_n N$, k and n . Then $2\ln(\tilde{g}_n)$ is asymptotically distributed as a chi-squared with one degree of freedom.*

Proof. We verify all assumptions of Theorem 3. Since $E(\frac{X_n}{n}) = p$ and $V(\frac{X_n}{n}) = \frac{p(1-p)(N+nK)}{n^2(N+k)}$, then $Prob(|\frac{X_n}{n} - p| > \varepsilon) \leq \frac{p(1-p)(N+nK)}{\varepsilon^2 n^2(N+k)}$. As $N = \beta n$, then $\frac{X_n}{n}$ converges in probability to p . Therefore $\frac{V(X_n)}{V(Y_n|X_n)}$ converges in probability to 1, being $V(Y_n|X_n) = \frac{\frac{X_n}{n}(1-\frac{X_n}{n})(N+nK)}{(N+k)}$. Set $E(X_n) = \mu_n$ and $V(X_n) = \sigma_n^2$. The density of $\frac{X_n - \mu_n}{\sigma_n}$ uniformly converges to the normal density on compact sets, as it is proved in the Appendix. Now we verify the last assumption of Theorem 3. Set $h_n(t; s) = p_{Y_n|ns}(ns + t\sqrt{V(Y_n|ns)}|ns)\sqrt{V(Y_n|ns)}$. Since $h_n(t; s)$ is the density of the standardized Polya's random variable with parameters N , sN , k and n , then $h_n(t; s)$ converges to $\varphi(t)$ for every $s \in (0; 1)$. Moreover, as proved in the Appendix, the convergence to normal density is uniform on compact sets also with respect to s . Therefore $h_n(0; \frac{X_n}{n}) = p_{Y_n|X_n}(X_n|X_n)\sqrt{V(Y_n|X_n)}$ converges in probability to $\varphi(0)$. Hence all the assumptions of Theorem 3 are verified and the thesis follows. \square

Note that both $\hat{\theta}_n$ and the maximum likelihood estimate of θ converge in probability to θ (see Corollary 6 in the Appendix). Hence a test based on the likelihood ratio of Proposition 4 may replace the maximum likelihood ratio test.

Acknowledgements

The authors are very grateful to the referees for valuable comments which have led to improvement in the presentation.

Appendix

Normal and Polya's distribution. Polya's urn model and its extensions are widely studied in literature also from the asymptotic point of view (see, for example, Bai and Hu, 2005, and their references). Polya's distribution is a Generalized Hypergeometric distribution studied by Kemp and Kemp (1956). Hida and Akahira (2003) provided a normal approximation of some type of Generalized Hypergeometric distributions, which do not include Polya's distribution.

Here we give another asymptotic result: we prove the convergence of the Polya's density to normal density, underlining that this convergence is uniform on compact sets with respect to two parameters.

Theorem 5 *Given $\alpha, p \in (0, 1)$ and $k = \{0, 1, 2, \dots\}$, consider Polya's probability function:*

$$p_{X_N}(x) = \binom{n}{x} \frac{B(x + \frac{M}{k}, n - x + \frac{N-M}{k})}{B(\frac{M}{k}, \frac{N-M}{k})},$$

where $n = \lfloor \alpha N \rfloor$, $M = \lfloor pN \rfloor$. Then, given two sequences of real numbers z_N converging to z and p_N converging to $p \in (0, 1)$, the following limit holds:

$$\lim_{N \rightarrow \infty} p_{X_N}(\lfloor \mu_N + z_N \sigma_N \rfloor) \sigma_N = \varphi(z), \text{ where } \mu_N = \alpha p_N N, \sigma_N = \sqrt{\alpha p_N q_N N \frac{N + \alpha k N}{N + k}}$$

and $q_N = 1 - p_N$. Therefore the convergence is uniform on compact sets both in z and in p .

Proof. Set:

$$f_N(x) = \frac{\Gamma(n+1)\Gamma(\frac{M}{k}+x)\Gamma(\frac{N-M}{k}+n-x)\Gamma(\frac{N}{k})}{\Gamma(x+1)\Gamma(n-x+1)\Gamma(\frac{M}{k})\Gamma(\frac{N-M}{k})\Gamma(\frac{N}{k}+n)}, \quad (3)$$

where Γ is the gamma function. Note that if x is a natural number $f_N(x) = p_{X_N}(x)$. Put $x = \mu_N + z_N\sigma_N$ and note that $x \rightarrow \infty$ as $N \rightarrow \infty$. Furthermore, for every sequence of real numbers z_N converging to z , $0 \leq x \leq n$, for N sufficiently large. Writing n and M as a function of N , it follows that every gamma argument in (3) tends to infinity as $N \rightarrow \infty$. Therefore for all gamma functions we can use the Stirling's formula $\Gamma(\nu + 1) \sim \sqrt{2\pi}\nu^{\nu+\frac{1}{2}}e^{-\nu}$, where the notation $u_\nu \sim v_\nu$ stands for $\lim_{\nu \rightarrow \infty} \frac{u_\nu}{v_\nu} = 1$. For example

$$\Gamma(n+1) = \Gamma([\alpha N] + 1) \sim \sqrt{2\pi}\alpha^{n+\frac{1}{2}}N^{n+\frac{1}{2}} \left(1 + \frac{[\alpha N] - \alpha N}{\alpha N}\right)^{n+\frac{1}{2}} e^{-n},$$

and

$$\Gamma(n-x+1) = \Gamma([\alpha N] - \alpha p_N N - z_N c \sqrt{N} + 1),$$

$$\Gamma(n-x+1) \sim \sqrt{2\pi}\alpha^{n-x+\frac{1}{2}}q_N^{n-x+\frac{1}{2}}N^{n-x+\frac{1}{2}} \left(1 - \frac{z_N c - \frac{[\alpha N] - \alpha N}{\sqrt{N}}}{\alpha q_N \sqrt{N}}\right)^{n-x+\frac{1}{2}} e^{-n+x},$$

where $c = \frac{\sigma_N}{\sqrt{N}}$. Replacing all gamma functions in (3) with the analogous expressions, we get:

$$f_N(x) \sim \frac{1}{\sqrt{2\pi}\sigma_N} \frac{(1+a_0)^{b_0}(1+a_1)^{b_1}(1-a_2)^{b_2}(1-a_3)^{b_3}}{(1+a_4)^{b_4}(1-a_5)^{b_5}(1-a_6)^{b_6}(1-a_7)^{b_7}(1-a_8)^{b_8}},$$

$$\text{where } a_0 = \frac{[\alpha N] - \alpha N}{\alpha N}, a_1 = \frac{z_N c - \frac{1}{\sqrt{N}} + \frac{[p_N] - p_N N}{k\sqrt{N}}}{\frac{p_N}{k}(1+\alpha k)\sqrt{N}}, a_2 = \frac{z_N c + \frac{1}{\sqrt{N}} + \frac{[p_N] - p_N N}{k\sqrt{N}} - \frac{[\alpha N] - \alpha N}{\sqrt{N}}}{\frac{q_N}{k}(1+\alpha k)\sqrt{N}},$$

$$a_3 = \frac{1}{\frac{N}{k}}, a_4 = \frac{z_N c}{\alpha p_N \sqrt{N}}, a_5 = \frac{z_N c - \frac{[\alpha N] - \alpha N}{\sqrt{N}}}{\alpha q_N \sqrt{N}}, a_6 = \frac{1 - \frac{[p_N] - p_N N}{k}}{\frac{p_N N}{k}}, a_7 = \frac{1 + \frac{[p_N] - p_N N}{k}}{\frac{q_N N}{k}},$$

$$a_8 = \frac{1 - (\lfloor \alpha N \rfloor - \alpha N)}{\frac{N(1+\alpha k)}{k}}, \text{ and } b_0 = \lfloor \alpha N \rfloor + \frac{1}{2}, b_1 = \frac{p_N N}{k}(1 + \alpha k) + z_N c \sqrt{N} + \frac{\lfloor p_N \rfloor - p_N N}{k} - \frac{1}{2},$$

$$b_2 = \frac{q_N N}{k}(1 + \alpha k) - z_N c \sqrt{N} - \frac{\lfloor p_N \rfloor - p_N N}{k} + (\lfloor \alpha N \rfloor - \alpha N) - \frac{1}{2}, b_3 = \frac{N}{k} - \frac{1}{2},$$

$$b_4 = \alpha p_N N + z_N c \sqrt{N} + \frac{1}{2}, b_5 = \alpha q_N N - z_N c \sqrt{N} + (\lfloor \alpha N \rfloor - \alpha N) + \frac{1}{2}, b_6 = \frac{p_N N}{k} + \frac{\lfloor p_N \rfloor - p_N N}{k} - \frac{1}{2},$$

$$b_7 = \frac{q_N N}{k} - \frac{\lfloor p_N \rfloor - p_N N}{k} - \frac{1}{2}, b_8 = \frac{N(1+\alpha k)}{k} + (\lfloor \alpha N \rfloor - \alpha N) - \frac{1}{2}.$$

As $N \rightarrow \infty$, p_N tends to p , c tends to $\sqrt{\alpha p q (1 + \alpha k)}$ and a_i tends to 0, for every $i = 0, 1, \dots, 8$. It is convenient to write $(1 + a_i)^{b_i} = \exp(b_i(\ln(1 + a_i)))$, for every $i = 0, 1, \dots, 8$. Use Taylor's expansion of $\ln(1 + a_i)$, for every i : for example

$$b_1(\ln(1 + a_1)) \sim \frac{z_N^2 c^2}{2 \frac{p_N}{k}(1+\alpha k)} + z_N c \sqrt{N} + \frac{\lfloor p_N \rfloor - p_N N}{k} - 1.$$

After tedious algebraical calculations, we obtain the simpler formula:

$$f_N(x) \sim \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left(-\frac{1}{2} \frac{z_N^2 c^2}{\alpha p_N q_N (1+\alpha k)}\right).$$

Now $f_N(x) \sim p_{X_N}(\lfloor x \rfloor)$, since $p_{X_N}(\lfloor x \rfloor) = f_N(x(1 + \frac{\lfloor x \rfloor - x}{x}))$ and $\frac{\lfloor x \rfloor - x}{x}$ tends to 0 as $N \rightarrow \infty$. Therefore:

$$\lim_{N \rightarrow \infty} p_{X_N}(\lfloor \mu_N + z_N \sigma_N \rfloor) \sigma_N = \lim_{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left(-\frac{1}{2} \frac{z_N^2 c^2}{\alpha p_N q_N (1+\alpha k)}\right) \sigma_N = \varphi(z). \square$$

Remark. In the previous theorem we have the same limit if we use the mode of X_N instead of μ_N .

Corollary 6 Let \hat{p}_N be the maximum likelihood estimate of p and $\hat{\theta}_N = \frac{X_N}{n}$, with $\alpha = \frac{n}{N}$. For $\alpha < \frac{1}{k}$, when $2 \leq X_N \leq n - 2$ then $|\hat{\theta}_N - \hat{p}_N| < \frac{1}{n}$. Therefore \hat{p}_N converges in probability to p .

Proof. Consider the Polya's probability function as a function of p :

$$p_{X_N}(x; p) = \binom{n}{x} \frac{B(x + \frac{pN}{k}, n - x + \frac{N - pN}{k})}{B(\frac{pN}{k}, \frac{N - pN}{k})} =$$

$$= \binom{n}{x} \frac{p(p + \frac{k}{N}) \cdots (p + (x-1)\frac{k}{N})(1-p)(1-p + \frac{k}{N}) \cdots (1-p + (n-x-1)\frac{k}{N})}{(1 + \frac{k}{N}) \cdots (1 + (n-1)\frac{k}{N})}.$$

$$\text{Set } d(p) = \frac{\partial \ln(p_{X_N}(x; p))}{\partial p} = \sum_{i=0}^{x-1} \frac{1}{p + i\frac{k}{N}} - \sum_{j=0}^{n-x-1} \frac{1}{1-p + j\frac{k}{N}}. \text{ If } 2 \leq x \leq n-2 \text{ and}$$

$\alpha < \frac{1}{k}$ then $\alpha k + 1 < x < n - 1 - \alpha k$. By using $\frac{1}{b+at} \leq \frac{1}{b+a[t]} \leq \frac{1}{b+a(t-1)}$, for

$$t \geq 0 \text{ and } 0 \leq a < b, d(\frac{x}{n} - \frac{1}{n}) \geq n \int_0^{x-1} \frac{1}{x-1+\alpha kt} dt - n \int_0^{n-x} \frac{1}{n-x+1+\alpha k(t-1)} dt \geq 0 \text{ and}$$

$$d(\frac{x}{n} + \frac{1}{n}) \leq n \int_0^x \frac{1}{x+1+\alpha k(t-1)} dt - n \int_0^{n-x-1} \frac{1}{n-x-1+\alpha kt} dt \leq 0. \text{ Hence, when } X_N = x,$$

the maximum likelihood $\hat{p}_N \in [\frac{x}{n} - \frac{1}{n}; \frac{x}{n} + \frac{1}{n}]$ and the first assertion follows.

For asymptotic normality of p_{X_N} the probability of $\{X_N < 2\} \cup \{X_N > n-2\}$

converges to 0. Therefore $|\hat{\theta}_N - \hat{p}_N| \rightarrow 0$ in probability. Since $\hat{\theta}_N$ converges in

probability to p , then \hat{p}_N converges in probability to p .

References

Antonelli S., Regoli G., 2005. On the Poisson-binomial relative error. *Statistics & Probability Letters* 71, 249-256.

Bai, Z. D. and Hu, F., 2005. Asymptotic in randomized urn models. *The Annals of Applied Probability*, vol. 15, 914-940.

Billingsley, P., 1968. *Convergence of Probability Measures*. John Wiley & Sons, New York.

Burr, I. W., 1973. Some approximate relations between terms of the hypergeometric, binomial and Poisson distributions. *Communications in Statistics*, 1, 297-301.

Feller, W., 1950. *An Introduction to Probability Theory and Its Applications*.

Vol. I, Vol. II, John Wiley & Sons, Inc., New York.

Hida, E. and Akahira, M., 2003. An approximation to the generalized hypergeometric distribution. *Statistical Papers*, 44, 483-497.

Johnson, N. L., Kotz, S., Kemp, A. W., 1992. *Univariate Discrete Distributions*. John Wiley & Sons, Inc., New York.

Kemp, C. D., Kemp, A. W., 1956. Generalized Hypergeometric Distributions, *J. Roy. Stat. Soc., (B)*, 18, 202-211.

Ord, J. K., 1968. Approximations to distribution functions which are hypergeometric series. *Biometrika*, 55, 243-248.

Sródka, T., 1963. On approximations of hypergeometric distribution. *Zeszyty Naukowe Politechniki Łódzkiej*, 10, 5-17.