Local limit theorem for large deviations and statistical box-tests

Igor Semaev

Department of Informatics, University of Bergen
Hoyteknologisenteret, N-5020 Bergen, Norway
e-mail: `igor@ii.uib.no`

Abstract: Let n particles be independently allocated into N boxes, where the l-th box appears with the probability a_l. Let μ_r be the number of boxes with exactly r particles and $\mu = [\mu_{r_1}, \ldots, \mu_{r_m}]$. Asymptotical behavior of such random variables as N tends to infinity was studied by many authors. It was previously known that if Na_l are all upper bounded and n/N is upper and lower bounded by positive constants, then μ tends in distribution to a multivariate normal low. A stronger statement, namely a large deviation local limit theorem for μ under the same condition, is here proved. Also all cumulants of μ are proved to be $O(N)$.

Then we study the hypothesis testing that the box distribution is uniform, denoted h, with a recently introduced box-test. Its statistic is a quadratic form in variables $\mu - \mathbf{E}\mu(h)$. For a wide area of non-uniform a_l, an asymptotical relation for the power of the quadratic and linear box-tests, the statistics of the latter are linear functions of μ, is proved. In particular, the quadratic test asymptotically is at least as powerful as any of the linear box-tests, including the well-known empty-box test if μ_0 is in μ.

AMS 2000 subject classifications: Primary 62F03,60F10; secondary 62H15,60B12.
Keywords and phrases: random allocations, large deviations, box-test power.

1. Introduction

1.1. Prior Work

Let n particles be independently allocated into N boxes(cells), where the l-th box appears with the probability a_l. We denote $a = (a_1, \ldots, a_N)$. Let $\mu_r(a)$ be the number of boxes with exactly r particles and $\mu(a) = [\mu_{r_1}(a), \ldots, \mu_{r_m}(a)]$ for some $r_1, \ldots, r_m$. Asymptotical behavior of such random variables was studied by Mises, Okamoto, Weiss, Renyi, Békéssy, Kolchin, Sevast'ynov, Chistyakov and others; see [10] for the bibliography before 1975, also [9] and [13] for later developments and applications. Poisson, normal and multivariate normal distributions are the most common limit distributions for μ under different conditions as N tends to infinity. Okamoto [16] and Weiss [23] were first who proved that if the cell distribution is uniform and $\alpha = n/N$ is a positive constant, then μ_0 has asymptotically normal distribution. The fact got a number of generalizations since then. In particular, Renyi [18] extended the area of α, where μ_0 is still asymptotically normal. The normality of μ_0 in case of the multinomial cell distribution a, where Na_l are upper bounded and α is upper and lower bounded by positive constants, was proved by Chistyakov [2]; see global and local limit Theorems 3.4.2 and 3.4.3 in [10]. In somewhat weaker condition that was also proved in [8]. A multidimensional normal theorem for μ was established in [22] under the same conditions as in [2]; see Theorems 3.5.2 and 3.5.3 in [10].

Those limit theorems were mostly produced with either the method of moments or characteristic functions. Both imply the convergence in distribution [11]. The latter is quite restrictive when it comes to studying the power of statistical box-tests, where it is often required to estimate the probability of an interval which may vary as N grows and that probability tends to zero rapidly. Large deviation limit theorems are then necessary. Such a theorem is one of the goals of the present work.

Any statistical test based on the distribution of μ is naturally to call a box-test. For instance, a test based on μ_0 is called empty box(cell)-test and was introduced by David in [6] to verify whether a sample of n independent observations is taken from a continuous distribution $F(x)$. The real axis is split into N subintervals $(z_{l-1}, z_l]$ of equal probability: $F(z_l) - F(z_{l-1}) = 1/N$. Then μ_0 is the number of intervals without observations. The test may have some advantage over Pearson's χ^2 goodness-of-fit test, which originally requires N fixed while $n \to \infty$ to approach limit distribution; see [24]. As the number of the intervals is fixed, some information on F may be lost. Though in practice χ^2-test is applicable for N growing as $n^{2/5}$; see [12], one can take advantage of a box-test by choosing the number of intervals N as big as n, for instance.

More general is a linear box-test whose statistic is the dot-product μc, where c is a constant non-zero vector of length m, we use the same character to denote a vector and its transpose. The power of the linear test was studied in detail. Let α be upper and lower bounded by positive constants, $N^{\frac{3}{2}} \sum_{l=1}^{N} (a_l - 1/N)^2$ tend to a constant and some additional restrictions be fulfilled. Then c was constructed such that the μc-test overcomes in power any test based on the same μ; see Chapter 5 in [10].

1.2. New box-tests

Let $Na_l \le C$ for all $l = 1, \dots, N$, where $C > 1$ is a constant, and let $\alpha = \frac{n}{N}$ be upper and lower bounded by positive constants: $0 < \alpha_0 \le \alpha \le \alpha_1$ as N and n tend to infinity. We assume that throughout the article: all statements below are true within this assumption. The random variable

$$\nu(a) = \frac{\mu(a) - \mathbf{E}\mu(a)}{\sqrt{N}} = \left(\frac{\mu_{r_1}(a) - \mathbf{E}\mu_{r_1}(a)}{\sqrt{N}}, \dots, \frac{\mu_{r_m}(a) - \mathbf{E}\mu_{r_m}(a)}{\sqrt{N}}\right)$$

asymptotically has then multivariate normal distribution; see Theorem 3.5.2 in [10]. This basic fact is employed to test whether a multinomial sample was produced with prescribed box probabilities for large enough N. The hypothesis $a = h$, where $h = (1/N, \dots, 1/N)$, is examined.

Let $\eta(a) = [\mu(a) - \mathbf{E}\mu(h)]N^{-1/2}$, where a is unknown box distribution. A quadratic box-test, whose statistic is the quadratic form $\eta\mathbf{B}^{-1}\eta$, where $\mathbf{B}$ is the covariance matrix for $\nu = \eta(h) = \nu(h)$, is here studied. Asymptotically, $\nu\mathbf{B}^{-1}\nu$ has χ^2_m-distribution as N tends to infinity. The limit covariance matrix defined by (8.1) in Appendix may be taken for $\mathbf{B}$. The test was found by this author and Hassanzadeh during a study on cryptographic hash-functions [21]. A good hash-function should have values indistinguishable from those produced with multinomial uniform probabilities. Hash-function values are considered as allocations into boxes, labeled with its different values. According to NIST, the total number of a hash function different values may be up to 2^{512} [15]. Therefore, to test a hash-function its values are split into N regions of equal probability as with the continuous distribution earlier.

The goal of this work is to prove Theorems 1.1 and 1.2. They imply in case of $N^{\frac{3}{2}}\sum_{l=1}^{N}(a_l-1/N)^2 \to \infty$ that the quadratic box-test is typically more powerful compared to any linear box-test, whose statistic depends on the same μ. The latter includes empty-box test if μ depends on μ_0.

The main technical tool is a local limit Theorem 1.3 for large deviations of μ proved in Section 3. The Theorem is of independent interest as it is a significant improvement over the Okamoto-Weiss result and a number of limit Theorems stated in [10].

Also, as a Corollary to Theorem 3.1, we show that within the assumptions all cumulants of μ are $O(N)$. This was not proved before.

1.3. Results

Let

$$\delta = \left(\frac{\mathbf{E}\mu_{r_1}(a) - \mathbf{E}\mu_{r_1}(h)}{\sqrt{N}}, \dots, \frac{\mathbf{E}\mu_{r_m}(a) - \mathbf{E}\mu_{r_m}(h)}{\sqrt{N}}\right). \tag{1.1}$$

and $\beta_c(a), \beta(a)$ denote second error probabilities of the linear ηc-test and the quadratic $\eta\mathbf{B}^{-1}\eta$-test with the same significance level; see Section 2.2.

Theorem 1.1. *Let $|\delta| \to \infty$ such that $|\delta| = o(N^{1/2})$. If $|\delta c| \to \infty$, then*

$$\frac{\ln \beta_c(a)}{\ln \beta(a)} = \frac{|\delta c|^2}{(c\mathbf{A}c)(\delta\mathbf{A}^{-1}\delta)}\,(1 + o(1)), \tag{1.2}$$

where $\mathbf{A}$ is a matrix defined by Theorem 8.2 in Appendix. If $|\delta c|$ is bounded, then $\ln \beta_c(a) = o(\ln \beta(a))$.

The matrix $\mathbf{A}$ is positive definite so, by the Cauchy-Schwarz inequality, $|\delta c|^2 \le (c\mathbf{A}c)(\delta\mathbf{A}^{-1}\delta)$. That inequality is often strict; see examples in Sections 5.1 and 5.2. Anyway, $\beta_c(a) \ge \beta(a)^{1+o(1)}$ and the quadratic test is typically of greater power than any linear in the Theorem assumptions. The next statement defines a, where the latter are fulfilled. We say $Na_l \to 1$ if for any $\varepsilon > 0$ there exists N_ε such that $|Na_l - 1| < \varepsilon$ for all $N > N_\varepsilon$ and $l = 1, \dots, N$.

Theorem 1.2. *Let $Na_l \to 1$, then $|\delta| = o(N^{1/2})$. Let in addition $(\alpha - r_i)^2 \ne r_i$ for some i. Then $|\delta| \to \infty$ if and only if $N^{\frac{3}{2}}\sum_{l=1}^{N}(a_l - 1/N)^2 \to \infty$.*

Theorem 1.1 is true for $a_l = N^{-1} + \gamma_l N^{-5/4}$, where $|\gamma_l|$ all tend to infinity such that $\gamma_l = o(N^{1/4})$. The power of the quadratic box-test then tends to 1 by Theorem 4.1. In other words, the quadratic test can distinguish such non-uniform distributions with the probability tending to 1 and provides with an asymptotically smaller error probability than linear box-tests.

Let $k = (k_1, \dots, k_m)$ be an integer vector and $x = (k - Np)N^{-1/2}$. We here denoted $p(a) = [p_{r_1}(a), \dots, p_{r_m}(a)]$, where $p_r(a)$ are defined by Theorem 3.1.5 in [10](Theorem 8.1 in Appendix), so that $\mathbf{E}\mu = Np + O(1)$. We study the probability $\mathbf{Pr}(\mu = k)$, where $|x|$ may grow as N grows. By $|x|$ the euclidean length of x is denoted.

Numerous local and global limit theorems for μ were proved; see [10] for a good account. The value of $|x|$ is there always assumed bounded. Also Berry-Esseen type estimates like those in [5] and [17] are too rough to estimate $\mathbf{Pr}(\mu = k)$ for large enough $|x|$. However a large deviation limit theorem, that is when $|x|$ is allowed to grow with N, is necessary for Theorem 1.1. In Section 3 we prove the following statement, where a modification of the saddle-point method due to Richter, see [19, 20], was used.

Theorem 1.3. *Assume $|x| = o(\sqrt{N})$. Then uniformly in α, a_l*

$$\boldsymbol{Pr}(\mu = k) \quad = \quad \frac{e^{-\frac{x\boldsymbol{A}^{-1}x}{2} + N\lambda(\frac{x}{\sqrt{N}})}}{\sqrt{(2\pi N)^m \det \boldsymbol{A}}} \left(1 + O\left(\frac{|x|+1}{\sqrt{N}}\right)\right),$$

where $\lambda(\tau) = O(|\tau|^3)$ for all small real m-vectors τ. The latter bound is uniform in α, N, a_l. The matrix $\boldsymbol{A} = (\sigma_{rt})$ is defined by (8.3).

This is a significant improvement over local limit Theorem 3.4.3 for μ_0 and multivariate Theorem 3.5.3 in [10]. Stronger versions of the global limit theorems, stated in [10] in case $Na_l \le C$ and $\alpha_0 \le \alpha \le \alpha_1$, are now easy to deduce.

By small latin and greek characters we denote both complex and real numbers and m-dimensional complex and real points. We hope it is clear from the context what this or that character means. Anyway, $x, s, \tau, t, v, \delta, c, p$ always(or almost always) denote m-dimensional real points, k is an m-dimensional integer point, and $u = s + it$ is an m-dimensional complex point, while z is usually a complex number and θ, σ are reals.

2. Quadratic Test

Let m-variate non-singular normal distribution be given. It is well known ([11], Section 15.10) that the quadratic form in the exponent of its density function has χ^2-distribution with m degrees of freedom when the quadratic form variables are substituted by the multivariate normal vector entries. Therefore the random variable $\nu\mathbf{B}^{-1}\nu$ asymptotically has χ^2-distribution with m degrees of freedom. Let $0 < \varepsilon < 1$ be a required significance level(criterion first kind error probability). The quantile C_ε such that

$$\mathbf{Pr}(\chi^2_m \ge C_\varepsilon) = \varepsilon \tag{2.1}$$

is taken from the χ^2_m distribution. Let an allocation of n particles into N boxes be observed. The statistic $\eta\mathbf{B}^{-1}\eta$ is computed, where $\eta(a) = [\mu(a) - Np(h)]N^{-1/2}$ and a is unknown box distribution. The former has the same asymptotical distribution as $[\mu(a) - \mathbf{E}\mu(h)]N^{-1/2}$ and is denoted by $\eta(a)$ again.

If $\eta\mathbf{B}^{-1}\eta \leq C_\varepsilon$, then $a = h$ is accepted, otherwise rejected. The first error probability $\mathbf{Pr}(\nu\mathbf{B}^{-1}\nu \geq C_\varepsilon)$ is close to ε for large enough N. For low N that may differ from ε and is computed directly; see Section 7 and [21].

2.1. Example

Let $\mu = (\mu_0, \mu_1)$ and $\alpha = 1$. Then by (8.1),

$$\mathbf{B} = \frac{1}{e^2} \times \begin{pmatrix} e-2 & -1 \\ -1 & e-1 \end{pmatrix}, \quad \mathbf{B^{-1}} = \frac{e^2}{e^2 - 3e + 1} \times \begin{pmatrix} e-1 & 1 \\ 1 & e-2 \end{pmatrix}$$

and

$$\eta(a) = \left(\frac{\mu_0(a) - Ne^{-1}}{\sqrt{N}}, \frac{\mu_1(a) - Ne^{-1}}{\sqrt{N}} \right).$$

So

$$\eta\mathbf{B}^{-1}\eta = \frac{e^2}{(e^2 - 3e + 1)N} \times \begin{pmatrix} \mu_0 - Ne^{-1} \\ \mu_1 - Ne^{-1} \end{pmatrix}^t \begin{pmatrix} e-1 & 1 \\ 1 & e-2 \end{pmatrix} \begin{pmatrix} \mu_0 - Ne^{-1} \\ \mu_1 - Ne^{-1} \end{pmatrix}.$$

2.2. Second kind error probabilities

The criterion second error probability is $\beta(a) = \mathbf{Pr}(\eta\mathbf{B}^{-1}\eta \leq C_\epsilon)$ for non-uniform a. In other words, that is the probability to accept $a = h$ whereas a is not uniform. Remark that $\beta(h) = 1 - \varepsilon$. When it comes to compare two criterions, one is said more powerful if its second error probability is lower while the first error probability is the same for both.

The dot-product ηc is the linear statistic under a. The second error probability is $\beta_c(a) = \mathbf{Pr}(|\eta c| \leq D_\epsilon)$, where D_ε is the $\mathbf{N}(0, \sqrt{c\mathbf{B}c})$-quantile of significance level ε. We will base linear and quadratic criteria on the same μ and compare their power in Section 4 by proving Theorem 1.1. Some limit theorems for large deviations are then required.

3. Local Limit Theorem for Large Deviations

By the characteristic function argument, see Theorem 3.5.3 in [10], the distribution of $(\mu - Np)N^{-1/2}$ tends to the multivariate normal distribution with 0 expectations and covariance matrix $\mathbf{A}$ defined in Theorem 8.2. So uniformly for any measurable area G

$$\mathbf{Pr}\left(\frac{\mu - Np}{\sqrt{N}} \in G \right) = \frac{1}{\sqrt{(2\pi)^s \det(\mathbf{A})}} \int_G e^{-\frac{x\mathbf{A}^{-1}x}{2}}\, dx + o(1).$$

If G varies with N, the main term may tend to 0 faster than the uniform estimate for $o(1)$. That makes the probability estimate useless. Stronger statements, where

$$\mathbf{Pr}\left(\frac{\mu - Np}{\sqrt{N}} \in G\right) = \frac{1 + o(1)}{\sqrt{(2\pi)^s \det(\mathbf{A})}} \int_G e^{-\frac{x\mathbf{A}^{-1}x}{2}}\, dx, \tag{3.1}$$

are then necessary. They are called limit theorems for large deviations and well-known for the sums of independent(especially identically distributed) random variables; see for instance [4, 7, 19, 20]. In [20] the sums of independent identically distributed multi-dimensional random variables were treated. Such a representation for μ is unknown, so the result isn't directly applicable.

Local limit Theorem 1.3 for large deviations of μ is here proved. Global estimates like (3.1) are then not difficult corollaries. The proof will follow that in [20] with some modifications. First, we have to prove Theorem 3.1, where the moment generating function for μ is represented in a form suitable for a saddle-point estimation, as in [20], of the integral in (3.2).

Let $s = (s_1, \ldots, s_m)$ be a vector of real variables and

$$\mathbf{M}_N(s) = \mathbf{E}e^{s\mu} = \sum_k \mathbf{Pr}(\mu = k)\, e^{sk}$$

the moment generating function for μ, where $sk = s_1k_1 + \ldots + s_mk_m$. There are finite number of terms in the sum as μ may take only finite number of values k. $\mathbf{M}_N(s)$ is a function in s which also depends on N, $n = \alpha N$, and the probabilities a_l. One can write

$$\mathbf{Pr}(\mu = k) = \frac{1}{(2\pi)^m} \int_{-\bar{\pi}}^{\bar{\pi}} \mathbf{M}_N(s + it) e^{-(s+it)k} dt, \tag{3.2}$$

where $t = (t_1, \ldots, t_m)$, and the integral is over $|t_j| \le \pi$, so $\bar{\pi} = (\pi, \ldots, \pi)$. By Theorem 3.1.1 in [10](Theorem 8.3 in Appendix) and Cauchy Theorem,

$$\begin{aligned} \mathbf{E}e^{u\mu} &= \frac{n!}{2\pi i N^n} \oint \prod_{l=1}^{N} F_l(z, u) \frac{dz}{z^{n+1}}, \\ F_l(z, u) &= e^{Na_lz} + \sum_{j=1}^{m} \frac{(Na_lz)^{r_j}}{r_j!}(e^{u_j} - 1) \end{aligned}$$

for any complex $u = s + it$, where $u = (u_1, \ldots, u_m)$, and complex variable z. The integral is over any closed contour encircling the origin. So one can write

$$\mathbf{M}_N(u) = \frac{n!}{2\pi i N^n} \oint e^{Nf(z,u)}\, \frac{dz}{z}, \tag{3.3}$$

where

$$\begin{aligned} f(z, u) &= \frac{1}{N}\left(\ln \prod_{l=1}^{N} F_l(z, u) - n \ln z\right) \\ &= \frac{1}{N} \sum_{l=1}^{N} \ln F_l(z, u) - \alpha \ln z. \end{aligned}$$

We take the main branch of ln for this definition.

Lemma 1. *Let s and z_0 take values from a closed bounded set in $\mathbb{R}^m$ and from a closed bounded interval of positive reals respectively. Let for $\varepsilon > 0$ either*

1. $\varepsilon \le |t_j| \le \pi$ *for at least one j and $|\theta| \le \pi$, or*
2. $\varepsilon \le |\theta| \le \pi$ *and $t = 0$*

be true. Then there exists $q < 1$ such that

$$\left|\prod_{l=1}^{N} \frac{F_l(z_0 e^{i\theta}, s+it)}{F_l(z_0, s)}\right| < q^N. \tag{3.4}$$

Proof. Denote $z = z_0 e^{i\theta}$, then

$$\left|\frac{F_l(z, s+it)}{F_l(z_0, s)}\right| = \frac{\left|e^{Na_l z} - \sum_{j=1}^m \frac{(Na_l z)^{r_j}}{r_j!} + \sum_{j=1}^m \frac{(Na_l z)^{r_j}}{r_j!} e^{s_j + it_j}\right|}{e^{Na_l z_0} - \sum_{j=1}^m \frac{(Na_l z_0)^{r_j}}{r_j!} + \sum_{j=1}^m \frac{(Na_l z_0)^{r_j}}{r_j!} e^{s_j}}.$$

Assume $a_l \ne 0$. As z_0 is positive,

$$\left|e^{Na_l z} - \sum_{j=1}^m \frac{(Na_l z)^{r_j}}{r_j!}\right| \le e^{Na_l z_0} - \sum_{j=1}^m \frac{(Na_l z_0)^{r_j}}{r_j!}.$$

If $\theta \ne 0$, then the last inequality is strict and so $|F_l(z, s+it)| < F_l(z_0, s)$. If $\theta = 0$, then the former inequality becomes equality, while the latter remains strict if at least one $t_j \ne 0$.

Let $C_1 < 1$ be a positive constant and $C_1 \le Na_l \le C$. Then for either of two conditions we have $|F_l(z, s+it)| < F_l(z_0, s)$. So

$$\left|\frac{F_l(z, s+it)}{F_l(z_\alpha, s)}\right| \le q_1$$

for a positive $q_1 < 1$. This is true because the function is continuous and its variables t, θ, z_0, s and Na_l are in a closed bounded set in both cases. More than $N(1 - C_1)(C - C_1)^{-1}$ probabilities a_l satisfy $C_1 \le Na_l \le C$. For every such l the product term in (3.4) is bounded by q_1. The rest are ≤ 1 in absolute value. Therefore (3.4) is true, where

$$q = q_1^{(1-C_1)(C-C_1)^{-1}} < 1.$$

□

Lemma 2. *There exist two closed bounded regions: multidimensional $U \subseteq \mathbb{C}^m$ around $u = 0$ and one-dimensional $V \subseteq \mathbb{C}$ that does not include $z = 0$. The regions do not depend on α, N and a_l. There exists a unique function $z_\alpha(u)$ on U with values in V such that $f'_z(z_\alpha(u), u) = 0$. The function $z_\alpha(u)$ is analytic on U.*

Proof.

$$f'_z(z,u)=\frac{1}{N}\sum_{l=1}^{N}\frac{Na_l\left(e^{Na_lz}+\sum_{j=1}^m\frac{(Na_lz)^{r_j-1}}{(r_j-1)!}(e^{u_j}-1)\right)}{e^{Na_lz}+\sum_{j=1}^m\frac{(Na_lz)^{r_j}}{r_j!}(e^{u_j}-1)}-\frac{\alpha}{z}.$$

is an analytic function in complex variables $z\in\mathbb{C}$ and $u\in\mathbb{C}^m$ in a small region around its zero $(\alpha,0)$ and $f''_{z^2}(\alpha,0)=\alpha^{-1}$. It follows from the Weierstrass Preparation Theorem [1] there are two regions: multidimensional U and one-dimensional V, such that for any $u\in U$ there exists only one $z_\alpha(u)\in V$ and $f'_z(z_\alpha(u),u)=0$. The regions U and V may be taken the same for any α, N and a_l within the assumptions.

We will prove that. Let $V\subseteq\mathbb{C}$ be the closed region bounded by the circle $z=\alpha'+\rho e^{i\theta}$ of radius $(\alpha_1-\alpha_0)/2<\rho<(\alpha_0+\alpha_1)/2$ centered at $\alpha'=(\alpha_0+\alpha_1)/2$. There exists $\gamma>0$ such that for any $\alpha_0\le\alpha\le\alpha_1$ the modulus of $f'_z(z,0)=\frac{z-\alpha}{z}$ on the border of V has its minimal value $\ge\gamma$. We now write

$$|f'_z(z,u)-f'_z(z,0)|=\left|\frac{1}{N}\sum_{l=1}^{N}\frac{Na_l\sum_{j=1}^m\left(\frac{(Na_lz)^{r_j-1}}{(r_j-1)!}-\frac{(Na_lz)^{r_j}}{r_j!}\right)(e^{u_j}-1)}{e^{Na_lz}+\sum_{j=1}^m\frac{(Na_lz)^{r_j}}{r_j!}(e^{u_j}-1)}\right|.$$

Let $z=\alpha'+\rho e^{i\theta}$. As $Na_l\le C$, there exists $r>0$ such that

$$|f'_z(z,u)-f'_z(z,0)|<\gamma$$

for any $|u|\le r$, α, N,a_l and any point z on the border of V. Let U denote the region $|u|\le r$. By Rouché's Theorem, the functions $f'_z(z,u)$ and $f'_z(z,0)$ have the same number of zeros z inside V, that is just one, for any $u\in U$. This defines a function $z_\alpha=z_\alpha(u)$ on U, which takes values in V. The Implicit Function Theorem [1] then states that $z_\alpha(u)$ is analytic on U. □

Lemma 3. *Any order partial derivatives of $z_\alpha(u)$ and $f(z_\alpha(u),u)$ are uniformly bounded on U while α, N and a_l may change.*

Proof. By Lemma 2, the function $z_\alpha(u)$, therefore $f(z_\alpha(u),u)$, are analytic on U and their values are uniformly bounded. By the Cauchy estimates [14], their any order partial derivatives are uniformly bounded on U too. □

Theorem 3.1. 1. *There exists a small region $U\subseteq\mathbb{C}^m$ around $u=0$, the same for all α, N and a_l, where*

$$\boldsymbol{M}_N(u)=e^{N\boldsymbol{K}(u)}\,Y_1(u)\,Y_2(u),\tag{3.5}$$

and $\boldsymbol{K}(u)$,$Y_1(u),Y_2(u)$ are analytic functions on U.

2. $\boldsymbol{K}(u)=up+\frac{u\boldsymbol{A}u}{2}+O(|u|^3)$, $\quad Y_1(u)=1+O(|u|)$ *uniformly in α, N and a_l.*

3. $Y_2(u)=O(N^{1/2})$ *for complex $u\in U$, and* $Y_2(s)=1+O(N^{-1/2})$ *for real $s\in U$. Both estimates are uniform in α,a_l.*

Proof. First we get the presentation (3.5) for $\mathbf{M}_N(s)$, where s belongs to a small closed bounded region in $\mathbb{R}^m$ around 0. The expansion of $\prod_{l=1}^N F_l(z,s)$ in z has only positive coefficients. So by Lemma 2.2.2 in [10](Theorem 8.4 in Appendix), for any real point s and positive α the equation $f'_z(z,s) = 0$ has a unique positive root $z_\alpha = z_\alpha(s)$ and $\alpha = z_\alpha(0)$. Also, by Theorem 8.5 in Appendix, $f''_{z^2}(z_\alpha(s),s)$ is then positive and, in particular, $f''_{z^2}(\alpha,0) = \alpha^{-1}$.

By Lemma 2, for all small enough s the value $z_\alpha(s)$ belongs to a closed bounded interval of positive real numbers. One can take the interval that does not depend on α, N and a_l. Also it is not difficult to show the values of partial derivatives $f^{(r)}_{z^r}(z_\alpha(s),s)$ are there uniformly bounded and $f''_{z^2}(z_\alpha(s),s)$ is uniformly upper and lower bounded by positive constants.

We estimate the integral in (3.3) for real s with the saddle-point method. The integration path is $z = z_\alpha(s)\exp(i\theta)$, where $|\theta| \le \pi$. Then $\mathbf{M}_N(s) = I_1 + I_2$, where

$$\begin{aligned} I_1 &= \frac{n!}{2\pi i N^n}\int_{|\theta|\le\varepsilon} e^{Nf(z,s)}\,\frac{dz}{z}, \\ I_2 &= \frac{n!}{2\pi i N^n}\int_{\varepsilon\le|\theta|\le\pi} e^{Nf(z,s)}\,\frac{dz}{z} \\ &= \frac{n!\,e^{Nf(z_\alpha,s)}}{2\pi i N^n}\int_{\varepsilon\le|\theta|\le\pi}\Big(\frac{z_\alpha}{z}\Big)^n\prod_{l=1}^N\frac{F_l(z,s)}{F_l(z_\alpha,s)}\,\frac{dz}{z}. \end{aligned}$$

The value ε will be chosen when estimating I_1. We estimate I_2 first. By Lemma 1, $\left|\prod_{l=1}^N F_l(z,s)F_l(z_\alpha,s)^{-1}\right| < q^N$, where $q<1$. Therefore,

$$|I_2| \le \frac{n!\,e^{Nf(z_\alpha,s)}}{N^n}q^N.$$

To estimate the first integral we expand

$$f(z_\alpha e^{i\theta},s) = f(z_\alpha,s) - \frac{f''_{z^2}(z_\alpha,s)}{2}\,z_\alpha^2\,\theta^2 + O(|\theta|^3)$$

around $\theta = 0$. The remainder estimate is uniform in α, N and a_l, for all small enough s. By a standard argument, see, for instance, Lemma 2.2.3 in [10](Theorem 8.6 in Appendix),

$$\mathbf{M}_N(s) \;=\; \frac{n!\,e^{Nf(z_\alpha(s),s)}(1+O(N^{-1/2}))}{N^n\,[2\pi\,N\,z_\alpha^2\,f''_{z^2}(z_\alpha(s),s)]^{1/2}}, \qquad (3.6)$$

where $(1+O(N^{-1/2}))$ is uniform in α, a_l and all small enough s. We now consider $z_\alpha(u)$ for complex u. With a tedious calculation one expands

$$z_\alpha(u) = \alpha + \sum_{j=1}^m\left(\frac{1}{N}\sum_{l=1}^N(Na_l\alpha - r_j)\frac{(Na_l\alpha)^{r_j}}{r_j!}e^{-Na_l\alpha}\right)u_j + \dots \qquad (3.7)$$

and

$$f(z_\alpha(u),u) = -\alpha\ln\alpha + \alpha + up + \frac{u\mathbf{A}u}{2} + O(|u|^3)$$

at $u=0$ uniformly by Lemma 3. We write $\mathbf{M}_N(u) = e^{N\mathbf{K}(u)}\,Y_1(u)\,Y_2(u)$, where

$$\begin{aligned}
\mathbf{K}(u) &= \alpha\ln\alpha - \alpha + f(z_\alpha(u),u),\\
Y_1(u) &= \left[\frac{\alpha}{z_\alpha^2(u)\,f''_{z^2}(z_\alpha(u),u)}\right]^{1/2},\\
Y_2(u) &= \frac{\mathbf{M}_N(u)}{e^{N\mathbf{K}(u)}\,Y_1(u)}.
\end{aligned}$$

From (3.6), uniformly $Y_2(s) = 1 + O(N^{-1/2})$ for real $s \in U$. That implies (3.5) for real s.

By direct calculation, $Y_1(u) = 1 + O(|u|)$ uniformly in α, N and a_l. All three above functions are analytic on U and the values of $\mathbf{K}(u)$ are there uniformly bounded. That proves the first two statements of the Theorem.

We prove $|Y_2(u)| = O(N^{1/2})$ for complex $u \in U$. It is enough to show that uniformly $|\mathbf{M}_N(u)\exp(-N\mathbf{K}(u))| = O(N^{1/2})$. Let's define the integration path in (3.3) by $z = z_\alpha(u)\exp(i\theta)$ and $|\theta| \le \pi$. We get

$$\left|\frac{\mathbf{M}_N(u)}{e^{N\mathbf{K}(u)}}\right| \le \frac{n!}{N^n e^{N(\alpha\ln\alpha-\alpha)}}\,\max_\theta\left|\frac{e^{Nf(z,u)}}{e^{Nf(z_\alpha(u),u)}}\right|. \tag{3.8}$$

By Stirling approximation to $n!$,

$$\frac{n!}{N^n e^{N(\alpha\ln\alpha-\alpha)}} = O(N^{1/2}). \tag{3.9}$$

Let s be a small real point, so $z_\alpha(s)$ belongs to a closed interval of positive reals and $u = s+it$. Any order partial derivatives of $f(z_\alpha(u)e^{i\theta},u) - f(z_\alpha(u),u)$ are bounded and it is represented by a power series for all small complex $u \in U$ and real $|\theta| \le \pi$. That series is a superposition of the convergent power series for $z_\alpha(u)$, $f(z,u)$ and $e^{i\theta}$. Therefore for every small s, we deduce a power series expansion in all small t and $|\theta| \le \pi$. By direct calculation,

$$\left|\frac{e^{f(z,u)}}{e^{f(z_\alpha(u),u)}}\right| = e^{-\left[\frac{f''_{z^2}(z_\alpha(s),s)\,z_\alpha^2(s)}{2} + O(|t|^2+|t|\theta)\right]\theta^2},$$

where the reminder estimate is uniform in α, N, a_l. Also we remark that $f''_{z^2}(z_\alpha(s),s)\,z_\alpha^2(s) > 0$ and it is uniformly bounded from below for small s. Therefore for all small enough t

$$\left|\frac{e^{Nf(z,u)}}{e^{Nf(z_\alpha(u),u)}}\right| \le 1.$$

It now follows from (3.8) and (3.9) that $|\mathbf{M}_N(u)\exp(-N\mathbf{K}(u))| = O(N^{1/2})$ in a small region of $u = s+it$ around $u=0$. We denote that region U again. That finishes the proof. □

Corollary 1. *All cumulants of μ are $O(N)$.*

Proof. According to [11], the cumulants are coefficients in the power series expansion for $\ln \mathbf{M}_N(u)$. The statement follows from Theorem 3.1. □

Lemma 4. *Let $|t_j| \le \pi$, where $1 \le j \le m$ and $\varepsilon \le |t_j|$ for at least one j. Let s belong to a small closed region in $\mathbb{R}^m$ around 0. Then there exists a positive $q < 1$, such that*

$$|\boldsymbol{M}_N(s+it)| \le e^{N\boldsymbol{K}(s)}\, q^N,$$

for all large enough N.

Proof. From (3.3) we get $|\mathbf{M}_N(s+it)| =$

$$\frac{n!\, e^{Nf(z_\alpha(s),s)}}{2\pi N^n}\left|\oint \left(\frac{z_\alpha}{z}\right)^n \prod_{l=1}^{N} \frac{F_l(z,s+it)}{F_l(z_\alpha(s),s)}\,\frac{dz}{z}\right|, \tag{3.10}$$

where the integration path is $z = z_\alpha(s)\exp(i\theta)$ and $|\theta| \le \pi$. By Lemma 1,

$$\left|\prod_{l=1}^{N} \frac{F_l(z_\alpha(s)\,e^{i\theta}, s+it)}{F_l(z_\alpha(s),s)}\right| < q^N,$$

where $q < 1$. For small s we have $f(z_\alpha(s),s) = -\alpha\ln\alpha + \alpha + \mathbf{K}(s)$. We use Stirling approximation to $n!$ and take a slightly larger q. The Lemma follows. □

Lemma 5.

$$\boldsymbol{Pr}(\mu = k) \le \frac{n!\, e^{Nf(z_\alpha(s),s)-sk}}{N^n}$$

for any real $s \in \mathbb{R}^m$ and $\boldsymbol{Pr}(\mu = k) = O(N^{1/2}\exp\left(N[\boldsymbol{K}(s) - \frac{sk}{N}]\right))$ for all small enough s.

Proof. One sees

$$\left|\frac{F_l(z_\alpha(s)\,e^{i\theta}, s+it)}{F_l(z_\alpha(s),s)}\right| \le 1,$$

for any real s, $|t| \le \pi$ and $|\theta| \le \pi$. From (3.10) we get $|\mathbf{M}_N(s+it)| \le n!\,N^{-n}\exp(Nf(z_\alpha,s))$. The Lemma follows from (3.2). □

One now proves Theorem 1.3.

Proof. Let $\mathbf{K}'(u)$ be the gradient of $\mathbf{K}(u) = \alpha\ln\alpha - \alpha + f(z_\alpha(u),u)$. From Lemma 3, the entries of $\mathbf{K}'(s)$ are functions in real $s \in U$ whose any order partial derivatives are uniformly bounded in α, N and a_l. Therefore, by Theorem 3.1 uniformly $\mathbf{K}'(s) = p + s\mathbf{A} + O(|s|^2)$. We can write (3.2) as

$$\mathbf{Pr}(\mu = k) = \frac{1}{(2\pi)^m}\int e^{N[\mathbf{K}(s+it) - \frac{(s+it)k}{N}]}\,Y(s+it)\,dt, \tag{3.11}$$

where $Y(u) = Y_1(u)\,Y_2(u)$ and estimate the integral with the saddle-point method. The saddle-point equation $\mathbf{K}'(s) = kN^{-1}$ is equivalent to $s\mathbf{A} + O(|s|^2) = \tau$, where $\tau = xN^{-1/2} = o(1)$.

By Theorem 8.2 in Appendix, the determinant of $\mathbf{A}$ is bounded from below by a positive constant. Transformation Inversion Theorem [3] states that for any small enough τ there exists unique solution $s = s(\tau)$ to $\mathbf{K}'(s) = kN^{-1}$ in a small region $U_1 \subseteq U$ around $s = 0$. That region is taken the same for all α, N and a_l. Any order partial derivatives of $s(\tau)$ are uniformly bounded. So uniformly $s = \tau\mathbf{A}^{-1} + O(|\tau|^2)$.

For small complex $u = s + it$ we write $\mathbf{M}_N(u) = \exp(N\mathbf{K}(u))\,Y(u)$, where $\mathbf{K}(u)$ and $Y(u) = Y_1(u)\,Y_2(u)$ are analytic functions defined in Theorem 3.1. One takes ε such that the following two Taylor expansions at $t = 0$ are true for all $|t_j| \le \varepsilon$ and $s \in U_1$.

$$\mathbf{K}(s+it) - \frac{(s+it)k}{N} = \mathbf{K}(s) - \frac{sk}{N} - \frac{H_2(t)}{2!} - \frac{iH_3(t)}{3!} + \ldots,$$

where $H_r(t)$ are forms of degree r. For instance, $H_2(t) = \sum_{i,j} t_i t_j \mathbf{K}''_{s_i,s_j}(s)$.

$$\frac{Y(s+it)}{Y(s)} = 1 + iR_1(t) - \frac{R_2(t)}{2!} + \ldots,$$

where $R_r(t)$ are forms of degree r. Remark $Y(s) \neq 0$ for small enough s and all large N. By Theorem 3.1, the values of $\mathbf{K}(u)$ are uniformly bounded on U while α, N, a_l changes. Therefore, by the Cauchy estimates, the coefficients in $H_r(t)$ are uniformly bounded too. As $|Y(u)| = O(N^{1/2})$ uniformly on U, then the coefficients in $R_r(t)$ are uniformly $O(N^{1/2})$. We split the integral in (3.11):

$$\begin{aligned}\mathbf{Pr}(\mu = k) \quad &= \quad \frac{1}{(2\pi)^m}\int_{-\bar{\varepsilon}}^{\bar{\varepsilon}} e^{N[\mathbf{K}(s+it) - \frac{(s+it)k}{N}]}\,Y(s+it)\,dt \\ &+ \quad \frac{1}{(2\pi)^m}\int_{\text{some }|t_j|\ge\varepsilon} \mathbf{M}_N(s+it)e^{-(s+it)k}\,dt.\end{aligned}$$

Let I denote the first integral. By Lemma 4, the second integral absolute value is bounded by $\exp(N[\mathbf{K}(s) - \frac{sk}{N}])\,q^N$, where $q < 1$. We change variables $tN^{1/2} = v$, use the above expansions and represent

$$I = \frac{e^{N[\mathbf{K}(s)-\frac{sk}{N}]}Y(s)}{(2\pi\sqrt{N})^m}\int_{-\bar{\varepsilon}\sqrt{N}}^{\bar{\varepsilon}\sqrt{N}} (1 + \frac{iR_1(v)}{\sqrt{N}} - \frac{R_2(v)}{2!N} + \ldots)\,e^{-\frac{H_2(v)}{2}}e^{-\frac{iH_3(v)}{3!\sqrt{N}}}\ldots\,dv.$$

So

$$I = \frac{e^{N[\mathbf{K}(s)-\frac{sk}{N}]}Y(s)}{(2\pi\sqrt{N})^m}\int_{-\bar{\varepsilon}\sqrt{N}}^{\bar{\varepsilon}\sqrt{N}} e^{-\frac{H_2(v)}{2}}\,T(v)\,dv,$$

where

$$T(v) = 1 + \frac{1}{\sqrt{N}}\left(\frac{R_1(v)H_3(v)}{3!\sqrt{N}} - \frac{R_2(v)}{2\sqrt{N}}\right) + \frac{1}{N}(\ldots).$$

Only forms of even degree appear in $T(v)$, as the integral for odd degrees is zero. We integrate the series term-wise with infinite integration limits. That introduces an error of order $e^{-C_1 N}$ for a constant $C_1 > 0$. One now writes

$$I = \frac{e^{N[\mathbf{K}(s)-\frac{sk}{N}]}Y(s)}{(2\pi N)^{m/2}\sqrt{\det H_2}}(1+O(N^{-1/2})).$$

Therefore, as $Y(s) = 1 + O(N^{-1/2})$ and $\det H_2 = \det \mathbf{A} + O(|\tau|)$, we have

$$\mathbf{Pr}(\mu = k) = \frac{e^{N[\mathbf{K}(s)-\frac{sk}{N}]}}{[(2\pi N)^m \det \mathbf{A}]^{1/2}}(1+O(|\tau| + N^{-1/2})) + O(e^{N[\mathbf{K}(s)-\frac{sk}{N}]}\, q^N).$$

As $\tau = \frac{k-Np}{N}$, by Theorem 3.1,

$$\mathbf{K}(s) \quad - \quad \frac{sk}{N} = -s\tau + \frac{s\mathbf{A}s}{2} + O(|s|^3),$$

Taking into account $s = \tau\mathbf{A}^{-1} + O(|\tau|^2)$, we get for all small enough τ

$$\mathbf{K}(s) - \frac{sk}{N} = -\frac{\tau\mathbf{A}^{-1}\tau}{2} + \lambda(\tau),$$
$$\lambda(\tau) = O(|\tau|^3)$$

uniformly in α, N, a_l. That implies the Theorem. □

4. Quadratic and Linear Test Power

In this section we estimate the second error probability of the quadratic and linear tests by using Theorem 1.3. That will imply Theorem 1.1. As $\delta = N^{1/2}\,[p(a) - p(h)] + O(N^{-1/2})$, we use δ to denote $N^{1/2}\,[p(a)-p(h)]$ in the rest of the article.

Theorem 4.1. *Let $|\delta| \to \infty$ such that $|\delta| = o(N^{1/2})$. Then*

$$\beta(a) = e^{-\frac{\delta\mathbf{A^{-1}}\delta}{2}(1+o(1))}.$$

Proof. By definition, $\beta(a) = \mathbf{Pr}\left(\eta\mathbf{B^{-1}}\eta \le C_\epsilon\right) = \sum_k \mathbf{Pr}(\mu = k)$, where k runs over $(x_k+\delta)\mathbf{B^{-1}}(x_k+\delta) \le C_\epsilon$ and $x_k = (k-Np)N^{-1/2}$ for $p = p(a)$. The matrix $\mathbf{B^{-1}}$ is positive definite. That implies $|x_k + \delta| \le C_1$ for a constant C_1 and so $|x_k| = o(N^{1/2})$. By Theorem 1.3,

$$\mathbf{Pr}(\mu = k) = \frac{e^{\theta(x_k)}(1+o(1))}{\sqrt{(2\pi N)^m \det \mathbf{A}}}, \tag{4.1}$$

where

$$\theta(x) = -\frac{x\mathbf{A}^{-1}x}{2} + N\lambda(\frac{x}{\sqrt{N}}).$$

Let k run over the set of integer m-vectors with non-negative entries, which satisfy $(x_k+\delta)\mathbf{B^{-1}}(x_k+\delta)\le C_\epsilon$. We write $\sum_k e^{\theta(x_k)}$ with repeated integrals by repeatedly using the identity

$$\sum_{n=a}^{b} f(n)=\int_a^b f(y)dy+\int_a^b (y-\lfloor y\rfloor)f'(y)dy+f(a), \tag{4.2}$$

where n runs over integers from a to b and $f(y)$ is one-variable function with continuous derivative. So we represent $\beta(a)$ by an integral in case of $m=1$. For $m>1$ the repeated integrals are written as multiple; see [3]. Anyway, we get

$$\sum_k e^{\theta(x_k)}=\int e^{\theta(x_k)}dk+H=\int e^{\theta(x_k)}dk\,(1+o(1)),$$

where H is a finite sum of multiple integrals whose multiplicity is lower than m or exactly m and the integrand is bounded by the absolute value of a partial derivative of $e^{\theta(x_k)}$. All these integrals are $o(\int e^{\theta(x_k)}dk)$. We now take $x=x_k+\delta$ as the integration variables. So $N^{1/2}dx=dk$ and

$$\begin{aligned}\sum_k e^{\theta(x_k)} &= N^{m/2}\int_{x\mathbf{B^{-1}}x\le C_\epsilon} e^{\theta(x-\delta)}dx\,(1+o(1))= \\ =\quad & C_1N^{m/2}e^{\theta(y-\delta)}\,(1+o(1))=N^{m/2}e^{-\frac{\delta\mathbf{A^{-1}}\delta}{2}(1+o(1))},\end{aligned}$$

by the Mean Value Theorem, where y is a point within the integration limits and $C_1=\int_{x\mathbf{B^{-1}}x\le C_\epsilon}dx$. Therefore, $y=O(1)$ and $\theta(y-\delta)=-\frac{\delta\mathbf{A^{-1}}\delta}{2}(1+o(1))$. The Theorem now follows from (4.1) and the definition of $\beta(a)$. □

Theorem 4.2. *Let c be any fixed real m-vector and $|\delta|\to\infty$ such that $|\delta|=o(N^{1/2})$. If $|\delta c|\to\infty$, then*

$$\beta_c(a)=e^{-\frac{|\delta c|^2}{2c\mathbf{A}c}(1+o(1))}.$$

If $|\delta c|$ is bounded, then $\beta_c(a)$ is bounded from below by a positive constant.

Proof. By definition, $\beta_c(a)=\mathbf{Pr}(|\eta c|\le D_\epsilon)$. Asymptotically ηc has a normal distribution with the expectation δc and the variance $c\mathbf{A}c$. From Theorem 8.3 and the Cauchy-Schwarz inequality, $c\mathbf{A}c$ is upper and lower bounded by positive constants. If $|\delta c|$ is bounded, then $\beta_c(a)$ tends to the probability of a bounded interval of length $2D_\epsilon$. Therefore, $\beta_c(a)$ is lower bounded by a positive constant in this case.

We represent $\beta_c(a)=\sum_k \mathbf{Pr}(\mu=k)$, where the sum is over k such that $|x_kc+\delta c|\le D_\epsilon$. Denote $x_k=(k-Np)N^{-1/2}=(x_{k1},\ldots,x_{km})$. Let $|\delta c|\to\infty$, then

$$\beta_c(a)=\sum_{\text{all }|x_{kj}|\le fN^{1/2}}+\sum_{\text{some }|x_{kj}|>fN^{1/2}}, \tag{4.3}$$

where f is a function such that $|\delta c|^2 = f^4 N$. Obviously, f tends to 0 and $fN^{1/2}$ tends to infinity. The first sum is represented by a sum of integrals as in the above proof and equal to $I + o(I)$, where

$$I = \frac{1}{\sqrt{(2\pi)^m \det A}} \int e^{\theta(x)} dx.$$

The integral is over $|xc + \delta c| \le D$, and $|x_j| \le fN^{1/2}$, where $x = (x_1, \ldots, x_m)$. With a standard transform of the integral and by the Mean Value Theorem,

$$I = e^{-\frac{|\delta c|^2}{2c\mathbf{A}c}(1+o(1))}.$$

The second sum in (4.3) is at most

$$\sum_{j=1}^{m} \mathbf{Pr}(|\mu_{r_j} - Np_{r_j}| \ge fN) = O(\sqrt{N} e^{N(-sf+O(s^2))}),$$

by Lemma 5, where we used the expansion of $\mathbf{K}(s)$ at $s = 0$ for $m = 1$. Put $s = f^2$, so the second sum is $o(I)$ too. That proves the statement. □

Theorem 1.1 now follows.

5. Examples

Two different non-uniform distributions a tending to uniform as N tends to infinity are considered in this Section. The power of the statistics ηc and $\eta\mathbf{B^{-1}}\eta$ on them are compared. The statistics only depend on μ_0 and μ_1.

5.1.

Let $N_1 \le N$, and $a_l = N_1^{-1}$ for $l = 1, \ldots, N_1$, and $a_l = 0$ for $l = N_1 + 1, \ldots, N$. Assume $\theta = NN_1^{-1}$ tends to 1 as N tends to infinity. By Theorem 8.2 one then computes the matrix $\mathbf{A}$ and the vector δ:

$$\mathbf{A} = \frac{1}{\theta e^{2\theta}} \times \begin{pmatrix} e^\theta - 1 - \theta & -\theta^2 \\ -\theta^2 & \theta e^\theta - \theta^3 + \theta^2 - \theta \end{pmatrix}.$$

and

$$\delta = N^{1/2} \left(\frac{e^{-\theta} + \theta - 1 - \theta e^{-1}}{\theta}, \frac{\theta e^{-\theta} - \theta e^{-1}}{\theta} \right).$$

Remark that δ satisfies the condition of Theorem 1.1, that is $|\delta| = o(N^{1/2})$. For different c, that is for different linear criterions, we study the expansion of the main term in the right-hand side of (1.2). For $c = (c_1, 1)$ the maximum of the expansion main term is achieved at $c_1 = (e-2)^{-1}(e-1)^{-1} = 0.8102..$ and

$$\frac{|\delta c|^2}{(c\mathbf{A}c)(\delta\mathbf{A^{-1}}\delta)} = 0.4272 + O((\theta - 1)).$$

That implies $\beta_{(c_1,1)} \geq \beta^{0.4272+o(1)}$ for all c_1. For $c = (1,0)$ we get $1+O((\theta-1)^2)$. So $\beta_{(1,0)} = \beta^{1+o(1)}$ and the empty-box test here is asymptotically as powerful as the quadratic test. The reason for that is the vectors

$$\begin{aligned} c\mathbf{A} &= \frac{1}{e^2}(e-2,-1) + O(\theta-1), \\ \delta &= \frac{(\theta-1)N^{1/2}}{e}(e-2,-1) + O((\theta-1)^2 N^{1/2}) \end{aligned}$$

are tending to be collinear. However that is not true in the next example.

5.2.

Let $N = 2N_1$ and $a_l = N^{-1} + \theta N^{-1}$ for $l = 1, \ldots, N_1$ and $a_l = N^{-1} - \theta N^{-1}$ for $l = N_1 + 1, \ldots, N$. Assume θ tends to 0 as N tends to infinity, therefore a tends to be uniform. The covariance matrix

$$\mathbf{A} = \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{01} & \sigma_{11} \end{pmatrix}$$

and the expectation vector δ for η are computed by Theorem 8.2:

$$\begin{aligned} \sigma_{00} &= \frac{e^{-1-\theta} - e^{-2-2\theta} + e^{-1+\theta} - e^{-2+2\theta}}{2} + \frac{\left[(1+\theta)e^{-1-\theta} + (1-\theta)e^{-1+\theta}\right]^2}{4} \\ \sigma_{01} &= -\frac{(1+\theta)e^{-2-2\theta} + (1-\theta)e^{-2+2\theta}}{2} \\ &\quad -\frac{\left[(1+\theta)e^{-1-\theta} + (1-\theta)e^{-1+\theta}\right]\left[(1+\theta)\theta e^{-1-\theta} - (1-\theta)\theta e^{-1+\theta}\right]}{4}, \\ \sigma_{11} &= \frac{(1+\theta)e^{-1-\theta} + (1-\theta)e^{-1+\theta}}{2} - \frac{(1+\theta)^2 e^{-2-2\theta} + (1-\theta)^2 e^{-2+2\theta}}{2}, \\ &\quad -\frac{\left[(1+\theta)\theta e^{-1-\theta} - (1-\theta)\theta e^{-1+\theta}\right]^2}{4}, \end{aligned}$$

and

$$\delta = N^{1/2}\left(\frac{e^{-1-\theta} + e^{-1+\theta}}{2} - e^{-1}, \frac{(1+\theta)e^{-1-\theta} + (1-\theta)e^{-1+\theta}}{2} - e^{-1}\right).$$

We study the expansion of the main term in (1.2). For instance, for $c = (1,1)$ that equals $O(\theta^4)$. So $\beta_{(1,1)} = \beta^{o(1)}$ and $(1,1)$-test behaves very poorly on such a. Generally, for $c = (c_1, 1)$ the maximum of the expansion main term in (1.2) is only achieved at $c_1 = (e-2)/(3-e) = 2.549..$ and equals $1 + O(\theta^2)$. So $\beta_{(c_1,1)} = \beta^{1+o(1)}$ in this case. For $c = (1,0)$

$$\frac{|\delta c|^2}{(c\mathbf{A}c)(\delta \mathbf{A^{-1}}\delta)} = 0.7469 + O(\theta^2).$$

The empty-box test is asymptotically less powerful than the quadratic test and $\beta_{(1,0)} = \beta^{0.7469+o(1)}$.

6. Where Theorem 1.2 is Proved

Let $a_l = N^{-1} + b_l$. Denote $\delta_r = N^{1/2}\,[p_r(a) - p_r(h)]$ and $R_N = N^{\frac{3}{2}} \sum_{l=1}^{N} b_l^2$.

Theorem 6.1. *Let $Na_l \to 1$, then*
1. $|\delta_r| = o(N^{1/2})$,
2. $|\delta_r| = O(R_N)$,
3. *assume in addition $(\alpha - r)^2 \neq r$, then $R_N = O(|\delta_r|)$.*

Proof. By (8.2),

$$p_r(a) = \frac{1}{N} \sum_{l=1}^{N} \frac{(\alpha N a_l)^r}{r!} e^{-\alpha N a_l}.$$

Let $x_l = Nb_l$. We put $f(x) = (1+x)^r e^{-\alpha x}$. By Taylor expansion, $f(x) = f(0) + (r-\alpha)x + f''(\theta x)\,x^2/2$, where $0 \leq \theta \leq 1$ around $x = 0$. For all large enough N

$$\begin{aligned}
\delta_r = \sqrt{N}\,[p_r(a) - p_r(h)] &= \frac{\alpha^r e^{-\alpha}}{r!\sqrt{N}} \sum_{l=1}^{N} [f(x_l) - f(0)] \\
&= \frac{\alpha^r e^{-\alpha}}{r!\sqrt{N}} \sum_{l=1}^{N} \left[(r-\alpha)\,x_l + f''(\theta_l\,x_l)\,x_l^2/2\right] \\
&= \frac{\alpha^r e^{-\alpha}}{r!\sqrt{N}} \sum_{l=1}^{N} f''(\theta_l\,x_l)\,x_l^2/2,
\end{aligned}$$

where $0 \leq \theta_l \leq 1$ and $\sum_{l=1}^{N} x_l = 0$. That implies the first statement. There exist two constants c_1, c_2 such that $c_1 \leq f''(x) \leq c_2$ for all small x. Therefore,

$$\frac{\alpha^r e^{-\alpha} c_1}{2\,r!} \left(N^{\frac{3}{2}} \sum_{l=1}^{N} b_l^2 \right) \leq \sqrt{N}\,[p_r(a) - p_r(h)] \leq \frac{\alpha^r e^{-\alpha} c_2}{2\,r!} \left(N^{\frac{3}{2}} \sum_{l=1}^{N} b_l^2 \right).$$

That implies the second statement. One computes $f''(0) = (\alpha - r)^2 - r$. If $(\alpha - r)^2 \neq r$, then c_1, c_2 may be taken both positive or both negative. That implies the last statement. □

Theorem 1.2 now follows from Theorem 6.1.

7. Tables

Let C_ε be defined by (2.1). As $N \to \infty$, the quadratic test first error probability $\mathbf{Pr}(\nu \mathbf{B}^{-1} \nu \geq C_\varepsilon) \to \varepsilon$. In this Section we experimentally study the convergence rate. Exact values of $\mathbf{Pr}(\nu \mathbf{B}^{-1} \nu \geq C_\varepsilon)$ for some μ and moderate N are presented, where $\varepsilon = 0.01$ and 0.05. Table 1 contains data for $\mu = (\mu_0, \mu_1, \mu_2, \mu_3)$ and Table 2 for $\mu = (\mu_2, \mu_4, \mu_5)$, where $n = N$. The probabilities are computed with the method explained in [21], which contains more computations. Numerical results demonstrate that the convergence rate depends on m, r_i and may be slow. Therefore, exact error probabilities should be used when applying the test for such low parameters.

TABLE 1
$(\mu_0, \mu_1, \mu_2, \mu_3)$-*quadratic test*

ε, N	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}
0.05	0.0564	0.0403	0.0621	0.0579	0.0527	0.0515	0.0507	0.0503	0.0500
0.01	0.0200	0.0229	0.0178	0.0171	0.0153	0.0134	0.0120	0.0111	0.0105

TABLE 2
(μ_2, μ_4, μ_5)-*quadratic test*

ε, N	2^4	2^5	2^6	2^7	2^8	2^9
0.05	0.0449	0.0907	0.0376	0.0561	0.0510	0.0522
0.01	0.0412	0.0163	0.0181	0.0134	0.0142	0.0120

Acknowledgements

I am grateful to V.F. Kolchin, B.A. Sevast'ynov and V.P. Chistyakov, the authors of "Random Allocations" [10], the reading of which was very stimulating.

References

[1] BOCHNER, S. AND MARTIN, W. (1948). *Several complex variables.* Princeton.

[2] CHISTYAKOV, V.P.(1964). *On the calculation of the power of the test of empty boxes.* Theory of Probab. Appl. **9** 648–653.

[3] COURANT, R. AND JOHN, F. (1974). *Introduction to calculus and analysis* **2**. J. Wiley & Sons, New York.

[4] CRAMÉR, H. (1938). *Sur un nouveau théorème-limit de la théorie des probabilités.* Actual. Sci. et Ind. No. 736, Paris.

[5] ENGLUND, G. (1981). *A remainder term estimate for the normal approximation in classical occupancy.* Ann. Probab. **9** 684–692.

[6] DAVID, F.N. (1950). *Two combinatorial tests whether a sample has come from a given population.* Biometrika. **37** 97–110.

[7] FELLER W. (1970). *An Introduction to Probability Theory and Its Applications* **2**. 2nd ed. J. Wiley & Sons, New York.

[8] HWANG, H.-K. AND JANSON, S. (2008). *Local limit theorems for finite and infinite urn models.* Ann. Probab. **36** 992–1022.

[9] JOHNSON, N.L. AND KOTZ, S.(1977). *Urn models and their applications.* J. Wiley & Sons, New York.

[10] KOLCHIN, V.F., SEVAST'YNOV, B.A. AND CHISTYAKOV, V.P. (1978). *Random Allocations.* V.H. Winston & Sons, Washington, D.C.

[11] KENDALL, M.G. AND STUART, A. (1969).*The Advanced Theory of Statistics* **1**. 3rd ed. Ch. Griffin & Company Limited, London.

[12] KENDALL, M.G. AND STUART, A. (1973). *The Advanced Theory of Statistics* **2**. Ch. Griffin & Company Limited, London.

[13] KOTZ, S. AND BALAKRISHNAN, N. (1997). *Advances in urn models during the past two decades.* In *Advances in Combinatorial methods and applications to probability and statistics.* Birkhäuser. Boston.

[14] Markushevich, A.I. (1985). *Theory of functions of a complex variables.* 2nd ed. Chelsea Pub Co.
[15] http://csrc.nist.gov/groups/ST/hash/documents/FR_Notice_Nov07.pdf
[16] Okamoto, M. (1952). *On a non-parametric test.* Osaka J. Math. **4** 77–85.
[17] Quine, M.P. and Robinson, J. (1982). *A Berry-Esseen bound for an occupancy problem.* Ann. Probab. **10** 663–671.
[18] Renyi, A. (1962). *Three new proofs and generalization of a theorem of Irving Weiss.* Magy. tud. akad. Mat. kutató int.kőzl. **7** 203–214.
[19] Richter, W. (1957). *Local Limit Theorems for Large Deviations.* Theory Probab. Appl. **2** 206–220.
[20] Richter, W. (1958). *Multi-Dimensional Local Limit Theorems for Large Deviations,* Theory Probab. Appl. **3** 100–106.
[21] Semaev, I. and Hassanzadeh, M. (2011). *New statistical box-test and its power.* available at http://www.ii.uib.no/publikasjoner/textap/pdf/2011-397.pdf
[22] Viktorova, I.I. and Chistyakov, V.P. (1965). *Asymptotic normality in a problem of balls when the probabilities of falling into different boxes are different.* Theory Probab. Appl. **10** 149–154.
[23] Weiss, I. (1958). *Limiting distributions in some occupancy problems,* Ann. Math. Stat. **29** 878–884.
[24] Wilks, S.S. (1962). *Mathematical Statistics.* J. Wiley & Sons, New York.

8. Appendix

In this Section we collect known auxiliary results mostly from [10], Chapter 3 with the exception of Theorem 8.2, which was not proved there. In Theorem 2.1.1 [10] it was proved that in case $a = h$

$$\mathbf{E}\mu_r = Np_r + O(1),$$
$$\mathbf{Cov}(\mu_r, \mu_t) = N\sigma_{rt} + O(1),$$

where

$$\begin{aligned} p_r &= \frac{\alpha^r}{r!}e^{-\alpha}, \\ \sigma_{rr} &= p_r - p_r^2 - p_r^2\frac{(\alpha - r)^2}{\alpha}, \\ \sigma_{rt} &= -p_r p_t - p_r p_t\frac{(\alpha - r)(\alpha - t)}{\alpha}. \end{aligned} \tag{8.1}$$

We put $\mathbf{B}(\alpha) = (\sigma_{r_i r_j})$. The matrix is positive definite and there exists a positive constant ρ such that $\det \mathbf{B}(\alpha) > \rho$ for all $\alpha_0 \le \alpha \le \alpha_1$ by Corollary 2.2.1 in [10]. More general are formulae (8.3) and Theorem 8.1 for box probabilities $a = (a_1, \dots, a_N)$. Let

$$p_{rl} = \frac{(\alpha N a_l)^r}{r!}e^{-\alpha N a_l}, \qquad p_r(a) = \frac{1}{N}\sum_{l=1}^{N} p_{rl}, \tag{8.2}$$

and

$$\sigma_{rr} = \frac{1}{N}\sum_{l=1}^{N} p_{rl} - \frac{1}{N}\sum_{l=1}^{N} p_{rl}^2 - \frac{1}{\alpha}\left[\frac{1}{N}\sum_{l=1}^{N} p_{rl}(\alpha N a_l - r)\right]^2, \quad (8.3)$$

$$\sigma_{rt} = -\frac{1}{N}\sum_{l=1}^{N} p_{rl}p_{tl} - \frac{1}{\alpha}\left[\frac{1}{N}\sum_{l=1}^{N} p_{rl}(\alpha N a_l - r)\right]\left[\frac{1}{N}\sum_{l=1}^{N} p_{tl}(\alpha N a_l - t)\right].$$

Theorem 8.1. (Theorem 3.1.5 in [10]) *Let N tend to infinity, $Na_l \le C$ for a constant C, and $\alpha_0 \le \alpha \le \alpha_1$. Then*

$$\mathbf{E}\mu_r(a) = Np_r + O(1),$$
$$\mathbf{Cov}(\mu_r(a), \mu_t(a)) = N\sigma_{rt} + O(1),$$

where p_r and σ_{rt} are defined by (8.2) *and* (8.3)*.*

Theorem 8.2. *Let $\mathbf{A} = (\sigma_{r_i r_j})$, where $1 \le i, j \le m$, as N tends to infinity such that $Na_l \le C$ for a constant C and $\alpha_0 \le \alpha \le \alpha_1$. Then $\mathbf{A}$ is positive definite and there exists a positive constant ρ such that* $\det \mathbf{A} \ge \rho$.

For $m = 1$ the statement is true by Theorem 3.1.4 in [10]. We here give a general proof.

Lemma 6. $x\mathbf{A}x \ge \frac{1}{N}\sum_{l=1}^{N} x\mathbf{B}(\alpha N a_l)\,x$ *for any real m-vector x, where $\mathbf{B}(\alpha N a_l)$ is defined by* (8.1)*.*

Proof. From (8.3)

$$\begin{aligned} x\mathbf{A}x &= \frac{1}{N}\sum_{l=1}^{N}\sum_{r} p_{rl}x_r^2 - \frac{1}{N}\sum_{l=1}^{N}\left(\sum_{r} p_{rl}x_r\right)^2 \\ &- \frac{1}{\alpha}\left(\frac{1}{N}\sum_{l=1}^{N}\sum_{r} p_{rl}(\alpha N a_l - r)x_r\right)^2, \end{aligned}$$

where r runs over $r_1, \dots, r_m$ and $x = (x_{r_1}, \dots, x_{r_m})$. Let $\gamma_l = \alpha N a_l$, then

$$\begin{aligned} \frac{1}{\alpha}\left(\frac{1}{N}\sum_{l=1}^{N}\sum_{r} p_{rl}(\alpha N a_l - r)x_r\right)^2 &= \alpha\left(\sum_{l=1}^{N} a_l \sum_{r}\frac{p_{rl}(\gamma_l - r)}{\gamma_l}x_r\right)^2 \\ \le \alpha\sum_{l=1}^{N} a_l \sum_{l=1}^{N} a_l\left(\sum_{r}\frac{p_{rl}(\gamma_l - r)}{\gamma_l}x_r\right)^2 &= \frac{1}{N}\sum_{l=1}^{N}\frac{1}{\gamma_l}\left(\sum_{r} p_{rl}(\gamma_l - r)x_r\right)^2 \end{aligned}$$

by the Cauchy-Schwarz inequality. Therefore,

$$\begin{aligned} x\mathbf{A}x &\ge \frac{1}{N}\sum_{l=1}^{N}\left[\sum_{r} p_{rl}x_r^2 - \left(\sum_{r} p_{rl}x_r\right)^2 - \frac{1}{\gamma_l}\left(\sum_{r} p_{rl}(\gamma_l - r)x_r\right)^2\right] \\ &= \frac{1}{N}\sum_{l=1}^{N} x\mathbf{B}(\gamma_l)\,x \end{aligned}$$

as $p_{rl} = \gamma_l^r e^{-\gamma_l}/r!$. □

So **A** is positive definite. We will prove Theorem 8.2 now.

Proof. Let $A_1, \ldots, A_m$ be the principal minor of $\mathbf{A}$. So $A_m = \det \mathbf{A}$. By Lemma 6, $A_1 \geq N^{-1} \sum_{l=1}^{N} \mathbf{B}_1(\gamma_l)$, where $\mathbf{B}_1(\gamma_l) = \sigma_{r_1 r_1}$ are principal minors of order 1 in matrices $\mathbf{B}(\gamma_l)$. Let $C_1 < 1$ be a positive constant. There are more than $N(1 - C_1)(C - C_1)^{-1}$ probabilities a_l such that $C_1 \leq N a_l \leq C$ and, therefore, $\alpha_0 C_1 \leq \gamma_l \leq \alpha_1 C$. By Corollary 2.2.1 in [10], there exists positive ρ_1 such that $\mathbf{B}_1(\gamma) > \rho_1$ for any $\alpha_0 C_1 \leq \gamma \leq \alpha_1 C$. So $A_1 \geq \rho_1(1 - C_1)(C - C_1)^{-1}$. Triangulating $x\mathbf{A}x$, one constructs $x = (x_1, \ldots, x_{m-1}, \sqrt{A_{m-1}})$, where $x\mathbf{A}x = A_m$. So

$$A_m \geq \frac{1}{N} \sum_{l=1}^{N} x\mathbf{B}(\gamma_l)x.$$

By Cauchy-Schwarz inequality, $(x\mathbf{B}(\gamma_l)x)(c\mathbf{B}(\gamma_l)^{-1}c) \geq |xc|^2$. We put $c = (0, \ldots, 0, 1)$ and get

$$A_m \geq \frac{A_{m-1}}{N} \sum_{l=1}^{N} \frac{1}{b_l},$$

where b_l is the bottom diagonal entry of $\mathbf{B}(\gamma_l)^{-1}$. For $C_1 \leq N a_l \leq C$ the entries of $\mathbf{B}(\gamma_l)$ are upper bounded and its determinant is lower bounded by a positive constant. Therefore, all entries of $\mathbf{B}(\gamma_l)^{-1}$ are upper bounded. In particular, $b_l < \rho_2$ for such l and some $\rho_2 > 0$. So $A_m \geq A_{m-1}(1 - C_1)\rho_2^{-1}(C - C_1)^{-1}$. The Theorem now follows by induction. □

Theorem 8.3. (Theorem 3.1.1 in [10]) *For any variables $z, x_1, \ldots, x_m$*

$$\sum_{n=0}^{\infty} \frac{(Nz)^n}{n!} \mathbf{E}\left[\prod_{j=1}^{m} x_j^{\mu_{r_j}(n,N)}\right] = \prod_{k=1}^{N}\left[e^{N a_l z} + \sum_{j=1}^{m} \frac{(N a_l z)^{r_j}}{r_j!}(x_j - 1)\right].$$

Let $F(z) = \sum_{k=0}^{\infty} a_l z^k$ be analytic function, where $a_l \geq 0$ for all k and $a_l > 0$ for all large enough k. Denote $f(z) = \ln F(z) - \alpha \ln z$.

Theorem 8.4. (Lemma 2.2.2 in [10]) *If $a_0 = a_1 = \ldots = a_{k_0 - 1} = 0$ and $a_{k_0} > 0$, then the equation $f'(z) = 0$ has a unique real positive root z_α for any $\alpha > k_0$.*

Theorem 8.5. *We have $f''(z_\alpha) > 0$ for any $\alpha > k_0$.*

Proof. For $z = z_\alpha$ one represents

$$f''(z) = \frac{zF''F - zF'^2 + F'F}{zF^2}.$$

Here $zF^2 > 0$ and $zF''F - zF'^2 + F'F = \sum_{l=0}^{\infty} b_l z^l$, where $b_l = \sum_{i=0}^{l} (l - i + 1)(l - 2i + 1) a_i a_{l-i+1} = \sum_{i=0}^{\lfloor \frac{l+1}{2} \rfloor} (l - 2i + 1)^2 a_i a_{l-i+1}$. Therefore $b_l > 0$ for all large enough l. That proves the statement.

Theorem 8.6. (Lemma 2.2.3 in [10]) *Suppose there exists a* $\gamma > 0$ *such that* $f''(z_\alpha) > \gamma$ *and* $z_\alpha > \gamma$*, where* $\alpha_0 \leq \alpha \leq \alpha_1$*. Then*

$$\frac{1}{2\pi i}\oint_{|z|=z_\alpha}\left[\frac{F(z)}{z^\alpha}\right]^N\frac{dz}{z} = \frac{e^{Nf(z_\alpha)}}{z_\alpha(2\pi f''(z_\alpha)N)^{1/2}}\left[1+O(N^{-1/2})\right].$$

Theorem 8.7. (Cauchy-Schwarz inequality) *Let* $\mathbf{A}$ *be a symmetric positive definite matrix of size* $m \times m$*. Then* $|xa|^2 \leq (x\mathbf{A}x)(a\mathbf{A^{-1}}a)$ *for any real* m*-vectors* x *and* a*.*