

On the Convergence of the Benjamini–Hochberg Procedure

Dean Palejev^{1,2,*}  and Mladen Savov^{1,3,*} 

¹ Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

² Big Data for Smart Society (GATE) Institute, Sofia University “St. Kliment Ohridski”, 1113 Sofia, Bulgaria

³ Faculty of Mathematics and Informatics, Sofia University “St. Kliment Ohridski”, 1164 Sofia, Bulgaria

* Correspondence: palejev@math.bas.bg (D.P.); mladensavov@math.bas.bg (M.S.)

Abstract: The Benjamini–Hochberg procedure is one of the most used scientific methods up to date. It is widely used in the field of genetics and other areas where the problem of multiple comparison arises frequently. In this paper we show that under fairly general assumptions for the distribution of the test statistic under the alternative hypothesis, when increasing the number of tests, the power of the Benjamini–Hochberg procedure has an exponential type of asymptotic convergence to a previously shown limit of the power. We give a theoretical lower bound for the probability that for a fixed number of tests the power is within a given interval around its limit together with a software routine that calculates these values. This result is important when planning costly experiments and estimating the achieved power after performing them.

Keywords: multiple hypothesis testing; Benjamini–Hochberg procedure; convergence rate



Citation: Palejev, D.; Savov, M. On the Convergence of the Benjamini–Hochberg Procedure. *Mathematics* **2021**, *9*, 2154. <https://doi.org/10.3390/math9172154>

Academic Editor: Christophe Chesneau

Received: 27 July 2021

Accepted: 31 August 2021

Published: 3 September 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In many modern studies scientists perform the same statistical test on many objects, e.g., genes or single-nucleotide polymorphisms. The number of these objects could be very large, say millions, resulting in a multiple comparison issue. This is due to the fact that even under the null hypothesis we expect some of these tests to have p -values below a predetermined significance level α . There are various techniques to overcome this problem with the Benjamini–Hochberg method [1] being the most used. Given α , for a set of m p -values, the method rejects the tests corresponding to the smallest k p -values, where k is the largest index such that $p_{(k)} \leq k\alpha/m$. The method controls the FDR (false discovery rate) in a sense that the expected proportion of false discoveries out of all discoveries (rejected hypotheses) is not more than α when the tests are independent as shown in the original paper [1]. In some cases, rather than finding the cutoff index for rejecting tests, the so-called Benjamini–Hochberg-adjusted p -values are used; these are the p -values transformed by the formula on the right-hand side of the above inequality and truncated at 1.

In the field of genetics it is used in many software packages, for example, DESeq2 [2], edgeR [3], cuffdiff [4], GenomeStudio (Illumina). It is also often used when studying various types of images, e.g., PET or CT [5], fMRI [6], or astrophysical [7].

Despite the fact that there are improved methods either in terms of weaker assumptions, e.g., dependency or better statistical power, e.g., [8–10], the Benjamini–Hochberg method is one of the most cited scientific articles, determined to be 59th most cited article in [11]. We should underline that in this work we are not interested in improving the Benjamini–Hochberg method, but rather than that in inferring about the statistical power of the original method as it is the most used one in practice. Moreover, article [12] shows that no correction of the standard methods is needed in certain cases even when there is some degree of dependence among the tests. See Section 5 for more information and related work.

Calculating the statistical power, that is the probability of correctly rejecting the null when the alternative hypothesis is true, is an important task when designing scientific

studies. For example it allows us to estimate the required sample size, and therefore the budget needed to perform the experiment, in order to make sure that we have a reasonable chance of detecting the objects for which there is indeed a difference between the two groups. After performing the study, it allows us to determine whether we have found a relatively large proportion of significant objects.

For practical applications we are interested not only in the limit of the statistical power as shown in previous works (see Section 5), but also in the speed of convergence to that limit. The speed of convergence would allow us to find the probability that the statistical power is within a given range around its limit. Consequently this would allow us to calculate the probability that the statistical power is above a pre-determined threshold, typically 80%. In cases when this probability is relatively low, one would have to increase the sample size in order to have more favorable parameters of the Beta distribution that models the p -values under the alternative hypothesis.

In more detail, in this work we show that under natural condition, that we call **(M)**, the convergence is exponential and give explicit expression of the constants involved in the asymptotics. That condition allows the usage of suitable Beta distributions amongst others. The relationship between our condition and the non-criticality condition in [13] will be discussed below. Condition **(M)** allows us to find a lower bound for the probability that for fixed parameters of the model the power is within given interval around that asymptotic limit. In this sense our work can be understood as a tool which can be used in the planning of the experiment and which complements [14], another article by the authors of [15], which deals with a variety of questions, such as estimation of the proportion of significant tests, the effect of the sample size to the quality of results, etc.

2. Main Results and Their Proofs

2.1. Some Preliminary Notation

Let $X_1, X_2, \dots, X_m, m \in \mathbb{N}^+$, with $\mathbb{N}^+ = \{1, 2, \dots\}$, be independent random variables that are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For some $1 \leq m_0 < m$ the random variables $(X_j)_{j \leq m_0}$ are assumed to have a common cumulative distribution function (c.d.f.) F , whereas $(X_j)_{m_0 < j \leq m}$ are assumed to have a common c.d.f. G . In this work we set F to be the c.d.f. of the uniform random variable on $(0, 1)$ and let for the time being G be any continuous c.d.f. We note that even if G is discrete, one can always approximate it with a suitable continuous distribution. We also work with the condition that $\frac{m_0}{m} = \gamma \in (0, 1)$ is fixed and set $m_1 := m - m_0 \in (0, m)$, which is strictly non-binding since we develop exponential bounds that are valid with different degree of accuracy for any fixed m_0, m and the respective ratio $\frac{m_0}{m}$. It will be usually the case that γ is larger than $1/2$ as the first m_0 random variables represent the non-significant observations. If $X_m^{(i)}, i \leq m$, is i -th order statistics of X_1, X_2, \dots, X_m then for a fixed level of rejection $\alpha \in (0, 1)$, the Benjamini–Hochberg procedure declares significant the first R_m order statistics, where

$$R_m = \max \left\{ i : X_m^{(i)} \leq \alpha \frac{i}{m} \right\}. \tag{1}$$

The truly non-significant tests have c.d.f F and the truly significant tests possess c.d.f. G . It is well known from [15] (1.2) that for $x \in (0, 1)$ we can express the event $\{R_m \geq mx\}$ in the following fashion

$$\left\{ \frac{R_m}{m} \geq x \right\} = \left\{ \max_{t \in A_\alpha(m, x)} \frac{H_m(t) - t}{t} \geq \frac{1}{\alpha} - 1 \right\}, \tag{2}$$

where $H_m(t) = \frac{1}{m} \sum_{j=1}^m \mathbb{I}_{\{X_j \leq t\}}$ is the empirical c.d.f. of X_1, X_2, \dots, X_m and

$$A_\alpha(m, x) = \left\{ \frac{k}{m} \alpha : \lceil mx \rceil \leq k \leq m \right\}, \tag{3}$$

where $\lceil y \rceil = \min\{m \in \mathbb{N}^+ : m \geq y\}$ is the ceiling function. We can express H_m as

$$H_m(t) = \frac{\gamma}{m_0} \sum_{j=1}^{m_0} \mathbb{I}_{\{X_j \leq t\}} + \frac{(1-\gamma)}{m_1} \sum_{j=m_0+1}^m \mathbb{I}_{\{X_j \leq t\}} = \gamma F_{m_0}(t) + (1-\gamma)G_{m_1}(t),$$

where F_{m_0} is the empirical c.d.f. of X_1, X_2, \dots, X_{m_0} and G_{m_1} is the empirical c.d.f. of X_{m_0+1}, \dots, X_m . Therefore a simple rearrangement yields that

$$\frac{H_m(t) - t}{t} = \gamma \frac{F_{m_0}(t) - t}{t} + (1-\gamma) \frac{G_{m_1}(t) - G(t)}{t} + (1-\gamma) \frac{G(t) - t}{t}. \tag{4}$$

2.2. General Inequalities and Information concerning the Distribution of R_m

We introduce the function $c(\epsilon) = (1 + \epsilon) \ln(1 + \epsilon) - \epsilon$, for $\epsilon > 0$, which is investigated in Proposition A1 in Appendix A. Then the following inequalities hold true.

Proposition 1. For any distribution function T with $T(\alpha x) > 0$ and empirical cumulative distribution function $T_n(\cdot) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}_{\{X_j \leq \cdot\}}$, where $(X_j)_{j \geq 1}$ are i.i.d. random variables with c.d.f. T , any $k, m \in \mathbb{N}^+$, any $\alpha, x, \rho \in (0, 1)$ and $A_\alpha(x, m)$ as in (3) above, we have, for every $\epsilon > 0$, that

$$\begin{aligned} & \mathbb{P}\left(\rho \max_{t \in A_\alpha(m, x)} \left| \frac{T_k(t) - T(t)}{t} \right| > \epsilon\right) \\ & \leq 2 \min \left\{ m(1-x)e^{-k \inf_{t \in A_\alpha(m, x)} T(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)}, e^{-2 \frac{k}{\rho^2} (\epsilon \alpha x)^2} \right\} \\ & := K_T(x, m, k, \rho, \epsilon, \alpha). \end{aligned} \tag{5}$$

Finally, if $T(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)$ as a function of t is non-decreasing on $[\alpha x, \alpha]$ we get that

$$\begin{aligned} & \mathbb{P}\left(\rho \max_{t \in A_\alpha(m, x)} \left| \frac{T_k(t) - T(t)}{t} \right| > \epsilon\right) \\ & \leq 2 \min \left\{ m(1-x)e^{-kT(\alpha x)c\left(\frac{\epsilon}{\rho} \frac{\alpha x}{T(\alpha x)}\right)}, e^{-2 \frac{k}{\rho^2} (\epsilon \alpha x)^2} \right\} = K_T(x, m, k, \rho, \epsilon, \alpha). \end{aligned} \tag{6}$$

Remark 1. Let us compute K_T when $T(t) = F(t) = t$ for $t \in (0, 1)$. Then $T(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right) = tc\left(\frac{\epsilon}{\rho}\right)$ is non-decreasing in t and substituting in the right-hand side of (6) we deduce that

$$K_F(x, m, k, \rho, \epsilon, \alpha) = 2 \min \left\{ m(1-x)e^{-k\alpha x c\left(\frac{\epsilon}{\rho}\right)}, e^{-2 \frac{k}{\rho^2} (\epsilon \alpha x)^2} \right\}. \tag{7}$$

Remark 2. Given that $c(\epsilon) \geq \min\left(1, \frac{\epsilon^2}{2e}\right)$, see Proposition A1 we may eradicate the dependence on the function c in K_T , see (5) by providing an upper bound for it:

$$\begin{aligned} & K_T(x, m, k, \rho, \epsilon, \alpha) \leq \\ & 2 \min \left\{ m(1-x)e^{-k \min\left(T(\alpha x), \inf_{t \in A_\alpha(m, x)} \frac{\epsilon^2}{\rho^2} \frac{t^2}{2eT(t)}\right)}, e^{-2 \frac{k}{\rho^2} (\epsilon \alpha x)^2} \right\}. \end{aligned} \tag{8}$$

Remark 3. Inspecting the proof below one easily discovers that for computational purposes it is better to use the sharper estimate

$$\mathbb{P}\left(\rho \max_{t \in A_\alpha(m, x)} \left| \frac{T_k(t) - T(t)}{t} \right| > \epsilon\right) \leq 2 \min \left\{ \sum_{t \in A_\alpha(m, x)} e^{-kT(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)}, e^{-2 \frac{k}{\rho^2} (\epsilon \alpha x)^2} \right\}. \tag{9}$$

For theoretical purposes the bounds stated in the theorem are more suitable.

Proof of Proposition 1. The second part of the minimum (5) follows from the Dvoretzky–Kiefer–Wolfowitz inequality, see [16], which gives, for any $x, \epsilon > 0$ that

$$\begin{aligned} \mathbb{P}\left(\rho \max_{t \in A_\alpha(m,x)} \left| \frac{T_k(t) - T(t)}{t} \right| > \epsilon\right) &\leq \mathbb{P}\left(\rho \sup_{t \in (\alpha x, \alpha)} \left| \frac{T_k(t) - T(t)}{t} \right| > \epsilon\right) \\ &\leq \mathbb{P}\left(\sup_{t \in (\alpha x, \alpha)} |T_k(t) - T(t)| > \frac{\epsilon \alpha x}{\rho}\right) \\ &\leq 2e^{-2k \frac{(\epsilon \alpha x)^2}{\rho^2}}. \end{aligned} \tag{10}$$

In Proposition A1 we use the function $c(\epsilon) = (1 + \epsilon) \ln(1 + \epsilon) - \epsilon$. The first part of the minimum of (5) is immediate from Proposition A1 since $kT_k(t) = \sum_{j=1}^k \mathbb{1}_{\{X_j \leq t\}}$ is binomially distributed with parameters k and $T(t) = \mathbb{P}(X_1 \leq t)$, and thus

$$\begin{aligned} \mathbb{P}\left(\rho \max_{t \in A_\alpha(m,x)} \left| \frac{T_k(t) - T(t)}{t} \right| > \epsilon\right) &\leq \sum_{t \in A_\alpha(m,x)} \mathbb{P}\left(\left|kT_k(t) - kT(t)\right| > \frac{\epsilon t}{\rho T(t)} kT(t)\right) \\ &\leq 2 \sum_{t \in A_\alpha(m,x)} e^{-kT(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)} \\ &\leq 2|A_\alpha(m,x)| e^{-k \inf_{t \in A_\alpha(m,x)} T(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)} \\ &\leq 2m(1-x) e^{-k \inf_{t \in A_\alpha(m,x)} T(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)}, \end{aligned}$$

where we have used the fact that the number of elements in $A_\alpha(m,x)$ is bounded from above as follows $|A_\alpha(m,x)| \leq m(1-x)$. Next, (6) is deduced from the assumption that $T(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)$ as a function of t is non-decreasing on $[\alpha x, \alpha)$ and the simple fact $A_\alpha(m,x) \subseteq [\alpha x, \alpha)$. \square

Applying Proposition 1 with $T = F$ and $T = G$ and respectively $m_0 = m\gamma$ and $m_1 = m(1-\gamma)$ and $\rho = \gamma$ and $\rho = 1-\gamma$ we can confirm from (4) the proof of Theorem 3.1 in [15] in the sense that we have the convergence in probability

$$\lim_{m \rightarrow \infty} \max_{t \in \left\{\alpha \frac{[mx]}{m}, \dots, \alpha \frac{m-1}{m}, \alpha\right\}} \frac{H_m(t) - t}{t} \stackrel{\mathbb{P}}{=} (1-\gamma) \sup_{[\alpha x, \alpha]} \frac{G(t) - t}{t} =: u_{\alpha, \gamma}(x). \tag{11}$$

Clearly, $u_{\alpha, \gamma} : [0, 1] \mapsto \mathbb{R}$, with $u_{\alpha, \gamma}(1) = \frac{G(\alpha) - \alpha}{\alpha}$, is a non-increasing function on $(0, 1]$. Denote

$$\underline{x}_{\alpha, \gamma}^\beta = \inf\{x \in (0, 1) : u_{\alpha, \gamma}(x) \leq \frac{1}{\beta} - 1\}, \quad \alpha, \beta \in (0, 1), \tag{12}$$

and

$$\bar{x}_{\alpha, \gamma}^\beta = \inf\{x \in (0, 1) : u_{\alpha, \gamma}(x) < \frac{1}{\beta} - 1\}, \quad \alpha, \beta \in (0, 1). \tag{13}$$

Since $u_{\alpha, \gamma}$ is non-increasing we have that $\bar{x}_{\alpha, \gamma}^\beta \geq \underline{x}_{\alpha, \gamma}^\beta$. Therefore from relations (2) and (4) we can easily get that

$$\begin{aligned} \lim_{m \rightarrow \infty} \mathbb{P}\left(\frac{R_m}{m} \geq x\right) &= 1, & \text{for } x < \underline{x}_{\alpha, \gamma}^\alpha \\ \lim_{m \rightarrow \infty} \mathbb{P}\left(\frac{R_m}{m} \geq x\right) &= 0, & \text{for } x > \bar{x}_{\alpha, \gamma}^\alpha. \end{aligned} \tag{14}$$

We can provide more precise information about probabilities of the type $\mathbb{P}\left(\frac{R_m}{m} < x\right)$ in the general case. In the sequel we use $\mathbb{I}_{\{A\}}$ for the indicator function of the set A .

Proposition 2. We have for any $\epsilon > 0$ and $x \in (0, 1)$ that

$$\mathbb{P}\left(\frac{R_m}{m} < x\right) \geq 1 - K_F(x, m, m_0, \gamma, \epsilon, \alpha) - K_G(x, m, m_1, 1 - \gamma, \epsilon, \alpha) - \mathbb{I}_{\left\{(1-\gamma) \max_{t \in A_\alpha(x, m)} \frac{G(t)-t}{t} \geq \frac{1}{\alpha} - 1 - 2\epsilon\right\}}. \tag{15}$$

Proof. Using successively (2), (4) and (10) with $\rho = \gamma, k = m_0 = \gamma m$ for the probability involving F_{m_0} and $\rho = 1 - \gamma, k = m_1 = (1 - \gamma)m$ for the expression concerning G_{m_1} , we arrive at the inequalities

$$\begin{aligned} \mathbb{P}\left(\frac{R_m}{m} < x\right) &= 1 - \mathbb{P}\left(\frac{R_m}{m} \geq x\right) \\ &= 1 - \mathbb{P}\left(\left\{\max_{t \in A_\alpha(x, m)} \frac{H_m(t) - t}{t} \geq \frac{1}{\alpha} - 1\right\}\right) \\ &\geq 1 - \mathbb{P}\left(\gamma \max_{t \in A_\alpha(x, m)} \left|\frac{F_{m_0}(t) - t}{t}\right| > \epsilon\right) - \mathbb{P}\left((1 - \gamma) \max_{t \in A_\alpha(x, m)} \left|\frac{G_{m_1}(t) - G(t)}{t}\right| > \epsilon\right) \\ &\quad - \mathbb{I}_{\left\{(1-\gamma) \max_{t \in A_\alpha(x, m)} \frac{G(t)-t}{t} \geq \frac{1}{\alpha} - 1 - 2\epsilon\right\}} \times \\ &\quad \times \mathbb{P}\left(\max_{t \in A_\alpha(x, m)} \left(\gamma \left|\frac{F_{m_0}(t) - t}{t}\right|, \left|(1 - \gamma) \frac{G_{m_1}(t) - G(t)}{t}\right|\right) \leq \epsilon\right) \\ &\geq 1 - K_F(x, m, m_0, \gamma, \epsilon, \alpha) - K_G(x, m, m_1, 1 - \gamma, \epsilon, \alpha) \\ &\quad - \mathbb{I}_{\left\{(1-\gamma) \max_{t \in A_\alpha(x, m)} \frac{G(t)-t}{t} \geq \frac{1}{\alpha} - 1 - 2\epsilon\right\}}. \end{aligned} \tag{16}$$

We only note that in the last expression of the first inequality above we have used that once the dependence on F_{m_0}, G_{m_1} is estimated away, the remaining term is deterministic and thus the indicator function appears. \square

2.3. Condition (M)

In view of the fact that Theorem 3.1 in [15] holds, the limit in probability, that is $\lim_{m \rightarrow \infty} \frac{R_m}{m} = x_{\alpha, \gamma}$ exists if and only if $\underline{x}_{\alpha, \gamma}^\alpha = \bar{x}_{\alpha, \gamma}^\alpha = x_{\alpha, \gamma}$. We shall therefore restrict our attention only to the case when for fixed α we have that the function $\frac{G(t)-t}{t}$ is decreasing on $(0, \alpha]$. In this case we have from (11) that

$$u_{\alpha, \gamma}(x) = (1 - \gamma) \frac{G(\alpha x) - \alpha x}{\alpha x}. \tag{17}$$

If in addition $\lim_{t \rightarrow 0} \frac{G(t)-t}{t} = \infty$ then

$$x_{\alpha, \gamma}^\beta := \underline{x}_{\alpha, \gamma}^\beta = \bar{x}_{\alpha, \gamma}^\beta, \text{ for any } \beta < 1, \tag{18}$$

and if $\beta = \alpha$ then we set $x_{\alpha, \gamma} := x_{\alpha, \gamma}^\alpha$. From now on we work under this condition which we call condition (M) and which is synthesized as follows.

Definition 1. We say that condition (M) holds for fixed $\alpha \in (0, 1)$ if and only if $H(t) := \frac{G(t)-t}{t}$ is decreasing on $(0, \alpha]$ with $\lim_{t \rightarrow 0} H(t) = \infty$.

Remark 4. As mentioned in [17], “[u]nder the alternative hypothesis, the p -values will have a distribution that has high density for small p -values and the density will decrease as the p -values

increase". Therefore when modeling the p -values under the alternative hypothesis with a Beta distribution, $B(a, b)$, its parameters should be so that $a < 1$ and $b \geq 1$. Note that condition **(M)** also holds when $b < 1$ and $a < (1 - 2\alpha)/(1 - \alpha)$, although that case should not happen for well-behaved statistical procedures.

Remark 5. Note that in our context Equation (2.2) of [13] translates to $\alpha_* = \inf_{u>0}\{u/G(u)\}$ and as it is mentioned there, the asymptotic of the Benjamini–Hochberg procedure exhibits a very different behavior on the regions $\alpha \in (0, \alpha_*)$ and $\alpha > \alpha_*$. In the same paper it is demonstrated that when $G(t)/t \rightarrow 0$ then the overwhelming proportion of discoveries will be amongst the non-significant features. Under our assumptions it is always true that $\alpha_* = 0$ and therefore we are always in the second regime, for which [13] provides a law of the iterated logarithm for R_n , see Theorem 2.2 of [13]. In more detail it is shown that the precise estimate

$$\limsup_{n \rightarrow \infty} \pm \frac{R_n - x_{\alpha,\gamma}}{\sqrt{n \log(\log(n))}} = \frac{\sqrt{x_{\alpha,\gamma}(1 - x_{\alpha,\gamma})}}{1 - \alpha G'(x_{\alpha,\gamma})}$$

holds true.

A very interesting contribution is made in Theorems 3.1 and 3.2 in [18]. The author finds a Donsker-type of convergence for the so-called threshold procedures, which are essentially functionals on the empirical functions $H_m(t)$ which recover, for example, the false discovery rate, see Theorem 3.2 in [18]. This implies a convergence rate $1/\sqrt{n}$ in $|R_n/n - x_{\alpha,\gamma}|$ and is applicable in a variety of other examples. Moreover, their results go beyond the scope of the Benjamini–Hochberg procedure and capture different procedures some of which are of higher power than the Benjamini–Hochberg procedure. In the particular case of the Benjamini–Hochberg procedure, however, we know from [15] that the probabilities of the events $\left\{ \left| \frac{R_n}{n} - x_{\alpha,\gamma} \right| \leq \beta x_{\alpha,\gamma} \right\}$ and of $\left\{ \left| \frac{S_n}{n} - x_{\alpha,\gamma} \frac{1-\alpha\gamma}{1-\gamma} \right| \leq \beta x_{\alpha,\gamma} \frac{1-\alpha\gamma}{1-\gamma} \right\}$, where S_n/n is the power of that procedure, converge to 1. Our Theorems 1 and 3 strengthen this by showing that the speed of convergence is of an exponential order. As discussed earlier, this allows a direct estimate on the probability that S_n/n is in a prescribed interval.

Remark 6. We note that upon the validity of condition **(M)** it is impossible that G has atoms on $(0, \alpha)$ as otherwise if $\mathbf{a} \in (0, \alpha)$ is an atom then

$$\lim_{t \downarrow \mathbf{a}} H(t) = H(\mathbf{a}) > \lim_{t \uparrow \mathbf{a}} H(t) = H(\mathbf{a}-),$$

which leads to a contradiction with the assumption that H is decreasing. Thus, our initial requirement that G is continuous is included in condition **(M)**. This also means that H is continuous on $(0, \alpha)$ and thus $\underline{x}_{\alpha,\gamma}^\beta = \bar{x}_{\alpha,\gamma}^\beta = x_{\alpha,\gamma}^\beta$.

We then have the following elementary lemma.

Lemma 1. Condition **(M)** is valid for fixed $\alpha \in (0, 1)$ if G has on $(0, \alpha)$ a density $g = G'$ satisfying $\lim_{x \rightarrow 0} g(x) = g(0) = \infty$ and $tg(t) - G(t) < 0$ on $(0, \alpha)$. The latter is satisfied if, for example, $g'(t) < 0$ on $(0, \alpha)$.

Proof. From the assumptions it is clear that on $(0, \alpha)$

$$\begin{aligned} \frac{dH(t)}{dt} &= \frac{(g(t) - 1)t - \int_0^t (g(s) - 1)ds}{t^2} < 0 \\ &\iff tg(t) - G(t) < 0. \end{aligned}$$

Furthermore, since $\lim_{t \rightarrow 0} \frac{G(t)}{t} = g(0) = \infty$ we conclude that $\lim_{t \rightarrow 0} H(t) = \infty$. Thus the first part of the proof is settled. Clearly, $tg(t) - G(t) < 0$ if $g'(t) < 0$ on $(0, \alpha)$ and the overall proof is thus completed. \square

The next lemma ensures an easier expression for K_G , see (5) and (6), provided condition (M) holds.

Lemma 2. *Let condition (M) hold for G and $m_1 = m(1 - \gamma)$. Then we have that*

$$K_G^M(x, m, 1 - \gamma, \epsilon, \alpha) = K_G(x, m, m_1, 1 - \gamma, \epsilon, \alpha) = 2 \min \left\{ m(1 - x)e^{-m(1-\gamma)G(\alpha x)c\left(\frac{\epsilon}{1-\gamma} \frac{\alpha x}{G(\alpha x)}\right)}, e^{-2\frac{m}{(1-\gamma)}(\epsilon \alpha x)^2} \right\}. \tag{19}$$

The same holds for the c.d.f. of the uniform distribution F . In particular if $\gamma > 1/2$ then

$$K_F^M(x, m, \gamma, \epsilon, \alpha) = K_F(x, m, m_0, \gamma, \epsilon, \alpha) = 2 \min \left\{ m(1 - x)e^{-m\gamma \alpha x c\left(\frac{\epsilon}{\gamma}\right)}, e^{-2\frac{m}{\gamma}(\epsilon \alpha x)^2} \right\}. \tag{20}$$

Remark 7. *Under condition (M) we have that $G(t)c\left(d \frac{t}{G(t)}\right)$ is non-decreasing on $(0, \alpha)$, for any $d > 0$, since $\frac{t}{G(t)}$ is non-decreasing on $(0, \alpha)$ and c is increasing on $(0, \infty)$, see Proposition A1. Hence the function $K_G^M(x, m, 1 - \gamma, \epsilon, \alpha)$ is decreasing in x as long as the other parameters stay fixed. The same is valid for K_F^M .*

Proof of Lemma 2. The proof is immediate from (6) since under condition (M) we have that the function $G(t)c\left(\frac{\epsilon}{1-\gamma} \frac{t}{G(t)}\right)$ is non-decreasing on $(0, \alpha)$, see Remark 7. We just note that by definition $m_1 = m(1 - \gamma)$ and we have employed this for k in (6) to derive (19). Finally, (20) follows from (7). \square

2.4. Theoretical Bounds on the Distribution of R_m under Condition (M)

If condition (M) holds we have the identity equivalent to (17)

$$\max_{t \in A_\alpha(x, m)} H(t) = \frac{G\left(\alpha \frac{[mx]}{m}\right) - \alpha \frac{[mx]}{m}}{\alpha \frac{[mx]}{m}} = H\left(\alpha \frac{[mx]}{m}\right) \tag{21}$$

and with $\beta = \alpha$ relation (18) together with (12) and (13) gives that

$$\frac{1}{\alpha} - 1 = (1 - \gamma)H(\alpha x_{\alpha, \gamma}) = (1 - \gamma) \left(\frac{G(\alpha x_{\alpha, \gamma}) - \alpha x_{\alpha, \gamma}}{\alpha x_{\alpha, \gamma}} \right). \tag{22}$$

Let $\beta > 0$ and denote using (21)

$$m_G(x, \beta, \alpha, \gamma) := \inf \{ k \in \mathbb{N}^+ : \max_{m \geq k} \mathbb{I} \{ (1-\gamma) \max_{t \in A_\alpha(x, m)} H(t) \geq \beta \} = 0 \} = \inf \{ k \in \mathbb{N}^+ : \max_{m \geq k} \mathbb{I} \{ (1-\gamma)H\left(\alpha \frac{[mx]}{m}\right) \geq \beta \} = 0 \}. \tag{23}$$

We then have the following key result:

Theorem 1. *Let condition (M) hold for given G and fixed $\alpha \in (0, 1)$ with $G(\alpha) < 1$. Fix $\gamma \in (0, 1)$. Recall from (18) that $x_{\alpha, \gamma} = \underline{x}_{\alpha, \gamma}^\alpha = \bar{x}_{\alpha, \gamma}^\alpha$. Then $\lim_{m \rightarrow \infty} \frac{R_m}{m} = x_{\alpha, \gamma} \in (0, 1)$. Moreover, for any $x \in (x_{\alpha, \gamma}, 1)$, $\exists \epsilon(x)$ such that for any $0 \leq \epsilon \leq \epsilon(x)$ the inequality $(1 - \gamma)H(\alpha x) < \frac{1}{\alpha} - 1 - 2\epsilon$ holds and then for any $m \geq m_G(x, \frac{1}{\alpha} - 1 - 2\epsilon, \alpha, \gamma)$ we have that*

$$\mathbb{P}\left(\frac{R_m}{m} \leq x\right) \geq 1 - K_F^M(x, m, \gamma, \epsilon, \alpha) - K_G^M(x, m, 1 - \gamma, \epsilon, \alpha). \tag{24}$$

Finally, for any $x \in (0, x_{\alpha,\gamma})$, $\exists \epsilon(x)$ such that for any $0 \leq \epsilon \leq \epsilon(x)$ the inequality $(1 - \gamma)H(\alpha x) > \frac{1}{\alpha} - 1 + 2\epsilon$ holds and then for any $m \geq m_G(x, \frac{1}{\alpha} - 1 + 2\epsilon, \alpha, \gamma)$

$$\mathbb{P}\left(\frac{R_m}{m} < x\right) \leq K_F^M(x, m, \gamma, \epsilon, \alpha) + K_G^M(x, m, 1 - \gamma, \epsilon, \alpha). \tag{25}$$

Remark 8. It seems important to evaluate or approximate $m_G(x, \frac{1}{\alpha} - 1 - 2\epsilon, \alpha, \gamma)$ with say m' so that (24) is valid beyond this estimate, that is for $m \geq m'$. If $H_d := \inf_{t \in (0, \alpha)} |H'(t)| > 0$, since $\alpha x < \alpha \frac{\lceil mx \rceil}{m}, 1 > x > x_{\alpha,\gamma}$ and H is decreasing on $(0, \alpha)$ we get using 22 in the first relation

$$\begin{aligned} \frac{1}{\alpha} - 1 &= (1 - \gamma)H(\alpha x_{\alpha,\gamma}) = (1 - \gamma)(H(\alpha x_{\alpha,\gamma}) - H(\alpha x)) + (1 - \gamma)H(\alpha x) \\ &\geq (1 - \gamma)H_d \alpha (x - x_{\alpha,\gamma}) + (1 - \gamma)H(\alpha x). \end{aligned}$$

However, if $(1 - \gamma)H_d \alpha (x - x_{\alpha,\gamma}) > 2\epsilon$ we get from the last observation that

$$\begin{aligned} (1 - \gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) &\leq (1 - \gamma)H(\alpha x) \leq \frac{1}{\alpha} - 1 - (1 - \gamma)H_d \alpha (x - x_{\alpha,\gamma}) \\ &< \frac{1}{\alpha} - 1 - 2\epsilon = \frac{1}{\alpha \epsilon} \end{aligned}$$

for any $m \geq 1$ and therefore if $(1 - \gamma)\alpha H_d (x - x_{\alpha,\gamma}) > 2\epsilon$ then

$$m_G(x, \frac{1}{\alpha} - 1 - 2\epsilon, \alpha, \gamma) = \inf\{k \in \mathbb{N} : \max_{m \geq k} \mathbb{I}_{\{(1-\gamma)H(\alpha \frac{\lceil mx \rceil}{m}) \geq \frac{1}{\alpha} - 1 - 2\epsilon\}} = 0\} = 1.$$

and (24) is always valid. Of course various similar and more precise estimates can be achieved, but these are beyond the scope of this work.

Remark 9. Considering Remark 3 the inequalities (24) and (25) can be improved for computational purposes by using as an upper bound the expression

$$2 \min \left\{ \sum_{t \in A_\alpha(m, x)} e^{-kT(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)}, e^{-2\frac{k}{\rho^2}(\epsilon \alpha x)^2} \right\}$$

for $K_T^M, T = G, F$ with $k = m_1, m_0; \rho = 1 - \gamma, \gamma$ respectively.

Remark 10. We point out that in our setting $\lim_{m \rightarrow \infty} \frac{R_m}{m} > 0$ and there have been a number of works that attempt to offer FDR even under weak dependence structure between the random variables following the uniform distribution or $(X_j)_{j \leq m_0}$. We point out, for example, to [19]. There is an interesting contribution, see [20], which discusses difficulties in controlling FDR under weak dependence in the case when $\lim_{m \rightarrow \infty} \frac{R_m}{m} = 0$.

Proof of Theorem 1. Under condition (M) we have that $x_{\alpha,\gamma}^\alpha = \bar{x}_{\alpha,\gamma}^\alpha := x_{\alpha,\gamma}$, see (18). Clearly, $x_{\alpha,\gamma} \in (0, 1)$ since $u_{\alpha,\gamma}(0) = \infty$ and $u_{\alpha,\gamma}(1) = \frac{G(\alpha) - \alpha}{\alpha} < \frac{1}{\alpha} - 1$, see (17), and by assumption $G(\alpha) < 1$. The fact that $\lim_{m \rightarrow \infty} \frac{R_m}{m} = x_{\alpha,\gamma} \in (0, 1)$ follows from Theorem 3.1 in [15]. Henceforth, it remains to show (24) and (25). The former thanks to (16) and (19) follows for any $m \geq m_G(x, \frac{1}{\alpha} - 1 - 2\epsilon, \alpha, \gamma)$, if $m_G(x, \frac{1}{\alpha} - 1 - 2\epsilon, \alpha, \gamma) < \infty$, since then

$$\mathbb{I}_{\{(1-\gamma) \max_{t \in A_\alpha(x, m)} \frac{G(t)-t}{t} \geq \frac{1}{\alpha} - 1 - 2\epsilon\}} = \mathbb{I}_{\{(1-\gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \geq \frac{1}{\alpha} - 1 - 2\epsilon\}} = 0.$$

However, since $\lim_{m \rightarrow \infty} \alpha \frac{\lceil mx \rceil}{m} = \alpha x$, $x > x_{\alpha, \gamma}$ and H decreasing on $(0, \alpha)$, then

$$\lim_{m \rightarrow \infty} (1 - \gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) = (1 - \gamma)H(\alpha x) < (1 - \gamma)H(\alpha x_{\alpha, \gamma}) = \frac{1}{\alpha} - 1.$$

Thus, for all ϵ small enough, we conclude that $(1 - \gamma)H(\alpha x) < \frac{1}{\alpha} - 1 - 2\epsilon$ and thus $m_G(x, \frac{1}{\alpha} - 1 - 2\epsilon, \alpha, \gamma) < \infty$, see (23). Hence, (24) follows.

To prove (25) for any $x \in (0, x_{\alpha, \gamma})$ we work with the complement event, that is $\left\{ \frac{R_m}{m} \geq x \right\}$. Similarly as in (16) and utilizing (4) and (21) we get that on the event

$$\left\{ \max_{t \in A_\alpha(x, m)} \left| \frac{F_{m_0}(t) - t}{t} \right| \leq \frac{\epsilon}{\gamma}; \max_{t \in A_\alpha(x, m)} \left| \frac{G_{m_1}(t) - G(t)}{t} \right| \leq \frac{\epsilon}{1 - \gamma} \right\}$$

we have that $\max_{t \in A_\alpha(x, m)} \frac{H_m(t) - t}{t} \geq (1 - \gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) - 2\epsilon$ and therefore

$$\begin{aligned} & \mathbb{P}\left(\frac{R_m}{m} \geq x\right) \geq \\ & \mathbb{P}\left(\frac{R_m}{m} \geq x; \max_{t \in A_\alpha(x, m)} \left| \frac{F_{m_0}(t) - t}{t} \right| \leq \frac{\epsilon}{\gamma}; \max_{t \in A_\alpha(x, m)} \left| \frac{G_{m_1}(t) - G(t)}{t} \right| \leq \frac{\epsilon}{1 - \gamma}\right) \\ & \geq \mathbb{P}\left(\max_{t \in A_\alpha(x, m)} \left| \frac{F_{m_0}(t) - t}{t} \right| \leq \frac{\epsilon}{\gamma}; \max_{t \in A_\alpha(x, m)} \left| \frac{G_{m_1}(t) - G(t)}{t} \right| \leq \frac{\epsilon}{1 - \gamma}\right) \\ & \times \mathbb{I}_{\left\{(1 - \gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \geq \frac{1}{\alpha} - 1 + 2\epsilon\right\}}. \end{aligned}$$

Again as $\frac{\lceil mx \rceil}{m} \rightarrow x < x_{\alpha, \gamma}$ and $H(\alpha x) > H(\alpha x_{\alpha, \gamma})$ we conclude that for all ϵ small enough the inequality $(1 - \gamma)H(\alpha x) \geq \frac{1}{\alpha} - 1 + 2\epsilon$ holds and we have that

$$\lim_{m \rightarrow \infty} \mathbb{I}_{\left\{(1 - \gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \geq \frac{1}{\alpha} - 1 + 2\epsilon\right\}} = 1. \tag{26}$$

From Proposition 1, the expressions for K_F^M, K_G^M in Lemma 2 and the independence of $(X_j)_{j \leq m_0}$ and $(X_j)_{m \geq j > m_0}$ we have that

$$\begin{aligned} & \mathbb{P}\left(\max_{t \in A_\alpha(x, m)} \left| \frac{F_{m_0}(t) - t}{t} \right| \leq \frac{\epsilon}{\gamma}; \max_{t \in A_\alpha(x, m)} \left| \frac{G_{m_1}(t) - G(t)}{t} \right| \leq \frac{\epsilon}{1 - \gamma}\right) \\ & = \mathbb{P}\left(\max_{t \in A_\alpha(x, m)} \left| \frac{F_{m_0}(t) - t}{t} \right| \leq \frac{\epsilon}{\gamma}\right) \mathbb{P}\left(\max_{t \in A_\alpha(x, m)} \left| \frac{G_{m_1}(t) - G(t)}{t} \right| \leq \frac{\epsilon}{1 - \gamma}\right) \\ & \geq \left(1 - K_F^M(x, m, \gamma, \epsilon, \alpha)\right) \left(1 - K_G^M(x, m, 1 - \gamma, \epsilon, \alpha)\right). \end{aligned}$$

and we conclude the proof of the theorem via the now trivial relations

$$\begin{aligned} & \mathbb{P}\left(\frac{R_m}{m} < x\right) = 1 - \mathbb{P}\left(\frac{R_m}{m} \geq x\right) \\ & \leq 1 - \mathbb{P}\left(\frac{R_m}{m} \geq x; \max_{t \in A_\alpha(x, m)} \left| \frac{F_{m_0}(t) - t}{t} \right| \leq \frac{\epsilon}{\gamma}; \max_{t \in A_\alpha(x, m)} \left| \frac{G_{m_1}(t) - G(t)}{t} \right| \leq \frac{\epsilon}{1 - \gamma}\right) \tag{27} \\ & \leq 1 - \left(1 - K_F^M(x, m, \gamma, \epsilon, \alpha)\right) \left(1 - K_G^M(x, m, 1 - \gamma, \epsilon, \alpha)\right) \times \\ & \times \mathbb{I}_{\left\{(1 - \gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \geq \frac{1}{\alpha} - 1 + 2\epsilon\right\}} \end{aligned}$$

and an appeal to (26). \square

For practical reasons and especially for the accompanying software routine it is optimal to choose ϵ in Theorem 1 as large as possible. The latter is due to fact that the bounds for K_F^M, K_G^M , see (19) and (20), are decreasing in ϵ . Since in practical computations m is usually

available, it is then beneficial to choose the largest ϵ as a function in x, m in a way that the previous theorem holds. In this direction we have the following result:

Theorem 2. Under the conditions of Theorem 1 and $m \in \mathbb{N}^+, x \in (x_{\alpha, \gamma}, 1)$ fixed we have that

$$\mathbb{P}\left(\frac{R_m}{m} \leq x\right) \geq 1 - K_F^M(x, m, \gamma, \epsilon_G, \alpha) - K_G^M(x, m, 1 - \gamma, \epsilon_G, \alpha), \tag{28}$$

where

$$\begin{aligned} \epsilon_G = \epsilon(x, m) &= \frac{1}{2}\left(\frac{1}{\alpha} - 1\right) - \frac{1 - \gamma}{2}H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \\ &= \frac{1 - \gamma}{2}\left(H(\alpha x_{\alpha, \gamma}) - H\left(\alpha \frac{\lceil mx \rceil}{m}\right)\right). \end{aligned} \tag{29}$$

Similarly, if $m \in \mathbb{N}, x \in (0, x_{\alpha, \gamma})$ are fixed and $\frac{\lceil mx \rceil}{m} < x_{\alpha, \gamma}$ then

$$\mathbb{P}\left(\frac{R_m}{m} < x\right) \leq K_F^M(x, m, \gamma, \epsilon_G, \alpha) + K_G^M(x, m, 1 - \gamma, \epsilon_G, \alpha), \tag{30}$$

where

$$\begin{aligned} \epsilon_G = \epsilon(x, m) &= \frac{1 - \gamma}{2}H\left(\alpha \frac{\lceil mx \rceil}{m}\right) - \frac{1}{2}\left(\frac{1}{\alpha} - 1\right) \\ &= \frac{1 - \gamma}{2}\left(H\left(\alpha \frac{\lceil mx \rceil}{m}\right) - H(\alpha x_{\alpha, \gamma})\right). \end{aligned} \tag{31}$$

Proof. To show (28) with ϵ_G as in (29) it suffices to show that for the chosen ϵ_G and the given m, x we have that

$$\mathbb{I}_{\{(1-\gamma) \max_{t \in A_\alpha(x, m)} \frac{G(t)-t}{t} \geq \frac{1}{\alpha} - 1 - 2\epsilon\}} = 0$$

in (16). However, as condition (M) holds, then

$$\max_{t \in A_\alpha(x, m)} \frac{G(t) - t}{t} = \max_{t \in A_\alpha(x, m)} H(t) = H\left(\alpha \frac{\lceil mx \rceil}{m}\right)$$

and we see that ϵ_G is the maximal possible choice such that

$$\mathbb{I}_{\{(1-\gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \geq \frac{1}{\alpha} - 1 - 2\epsilon\}} = 0.$$

For the second part of (29) we only recall that $(1 - \gamma)H(\alpha x_{\alpha, \gamma}) = \frac{1}{\alpha} - 1$, see (22). The rest of the theorem, that is (30) and (31), follows in a similar fashion from (27) wherein the expression of (31) is the maximal ϵ such that for the fixed $m, x \in (0, x_{\alpha, \gamma})$

$$\mathbb{I}_{\{(1-\gamma)H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \geq \frac{1}{\alpha} - 1 + 2\epsilon\}} = 1.$$

We only note that in this instance we need to have that $x < \frac{\lceil mx \rceil}{m} < x_{\alpha, \gamma}$, which is part of the assumptions of the theorem. \square

2.5. Exponential Convergence to the Theoretical Limit of the Power of the Benjamini–Hochberg Procedure

We proceed by studying the question about the power of the test. It is clear that if we denote by

$$S_m = \sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha \frac{R_m}{m}\}}, \tag{32}$$

where $(X_j)_{1 \leq j \leq m_1}$ are i.i.d. with c.d.f G then

$$S_m^* = \frac{S_m}{m_1} \tag{33}$$

is the proportion of tested elements that we have correctly identified as significant or the power of the test. If one wishes to understand the behavior of S_m^* with the increase of m one ought to decouple its dependence on $\frac{R_m}{m}$ in (32). Theorem 1 aids in this matter. We note that if we choose $x_1 < \lim_{m \rightarrow \infty} \frac{R_m}{m} = x_{\alpha, \gamma} < x_2$, see (18) for the definition of $x_{\alpha, \gamma}$, then from Theorem 1 we have for every

$$m \geq \max \left\{ m_G \left(x_1, \frac{1}{\alpha} - 1 - 2\epsilon, \alpha, \gamma \right); m_G \left(x_2, \frac{1}{\alpha} - 1 + 2\epsilon, \alpha, \gamma \right) \right\} := m^*(x_1, x_2, \epsilon, \alpha, \gamma)$$

with $\epsilon \leq \min \{ \epsilon(x_1), \epsilon(x_2) \}$ that

$$\begin{aligned} \mathbb{P} \left(\frac{R_m}{m} \in [x_1, x_2] \right) &= \mathbb{P} \left(\frac{R_m}{m} \leq x_2 \right) - \mathbb{P} \left(\frac{R_m}{m} < x_1 \right) \\ &\geq 1 - \left(K_F^M(x_2, m, \gamma, \epsilon, \alpha) + K_G^M(x_2, m, 1 - \gamma, \epsilon, \alpha) \right) \\ &\quad - \left(K_F^M(x_1, m, \gamma, \epsilon, \alpha) + K_G^M(x_1, m, 1 - \gamma, \epsilon, \alpha) \right). \end{aligned} \tag{34}$$

Then we can formulate the following fundamental result that offers exponential bounds on the convergence of S_m^* .

Theorem 3. *Let condition (M) be valid and assume the notation of Theorem 1. Let next $x_{\alpha, \gamma} \in (x_1, x_2) \subset (0, 1)$. Then there exists $m^* = m^*(x_1, x_2, \epsilon, \alpha, \gamma)$ and $\epsilon^* = \epsilon(x_1, x_2)$ such that for any $m \geq m^*$ and any $\epsilon \leq \epsilon^*$*

$$\begin{aligned} \mathbb{P} \left(\frac{1}{m_1} \sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x_1\}} \leq S_m^* \leq \frac{1}{m_1} \sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x_2\}} \right) \\ \geq 1 - \left(K_F^M(x_1, m, \gamma, \epsilon, \alpha) + K_G^M(x_1, m, 1 - \gamma, \epsilon, \alpha) \right) \\ - \left(K_F^M(x_2, m, \gamma, \epsilon, \alpha) + K_G^M(x_2, m, 1 - \gamma, \epsilon, \alpha) \right). \end{aligned} \tag{35}$$

Clearly, the random variables $\sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x\}}$ are Binomial with parameters m_1 and $G(\alpha x) = \mathbb{P}(X_1 \leq \alpha x)$. Henceforth,

$$\lim_{m \rightarrow \infty} S_m^* = G(\alpha x_{\alpha, \gamma}) = x_{\alpha, \gamma} \frac{1 - \alpha \gamma}{1 - \gamma}. \tag{36}$$

Moreover, for any $\eta \in (0, 1)$ with m, ϵ as above, we have that

$$\begin{aligned} \mathbb{P} \left((1 - \eta)G(\alpha x_1) \leq S_m^* \leq (1 + \eta)G(\alpha x_2) \right) \\ \geq 1 - \left(K_F^M(x_1, m, \gamma, \epsilon, \alpha) + K_G^M(x_1, m, 1 - \gamma, \epsilon, \alpha) \right) \\ - \left(K_F^M(x_2, m, \gamma, \epsilon, \alpha) + K_G^M(x_2, m, 1 - \gamma, \epsilon, \alpha) \right) \\ - e^{-mc(\eta)(1-\gamma)G(\alpha x_1)} - e^{-mc(\eta)(1-\gamma)G(\alpha x_2)} \end{aligned} \tag{37}$$

where $c(\eta) = (1 + \eta) \ln(1 + \eta) - \eta, \eta > 0$, is considered in Proposition A1.

Remark 11. We note that (36) has been proved in Corollary 3.3 in [15], but here we offer some exponential bounds on this convergence.

Proof of Theorem 3. The proof of (35) follows immediately from (32), (34) and the considerations thereabout since

$$\mathbb{P}\left(\frac{1}{m_1} \sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x_1\}} \leq S_m^* \leq \frac{1}{m_1} \sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x_2\}}\right) \geq \mathbb{P}\left(\frac{R_m}{m} \in [x_1, x_2]\right). \tag{38}$$

The fact that $\sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x\}}$ is Binomial with parameters m_1 and $G(\alpha x) = \mathbb{P}(X_1 \leq \alpha x)$ is straightforward which with the help of (35) and

$$\begin{aligned} &\lim_{m \rightarrow \infty} \left(K_F^M(x_1, m, \gamma, \epsilon, \alpha) + K_G^M(x_1, m, 1 - \gamma, \epsilon, \alpha) \right) \\ &+ \lim_{m \rightarrow \infty} \left(K_F^M(x_2, m, \gamma, \epsilon, \alpha) + K_G^M(x_2, m, 1 - \gamma, \epsilon, \alpha) \right) = 0 \end{aligned}$$

prove (36). Indeed, the strong law of large numbers gives almost surely that

$$\lim_{m_1 \rightarrow \infty} \frac{1}{m_1} \sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x\}} = G(\alpha x)$$

and from (35) $\lim_{m \rightarrow \infty} S_m^* = G(\alpha x_{\alpha, \gamma})$. Next, from (22) we get that

$$H(\alpha x_{\alpha, \gamma}) = \frac{G(\alpha x_{\alpha, \gamma}) - \alpha x_{\alpha, \gamma}}{\alpha x_{\alpha, \gamma}} = \frac{1 - \alpha}{\alpha}$$

and henceforth $G(\alpha x_{\alpha, \gamma}) = x_{\alpha, \gamma} \frac{1 - \alpha \gamma}{1 - \gamma}$. The latter is the limit on the right-hand side of (36). The very last relation, that is (37), follows again from (35) and an application of Proposition A1 which yields in our case with $m_1 = m(1 - \gamma)$ the bounds

$$\mathbb{P}\left(\frac{1}{m_1} \sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x_2\}} \geq (1 + \eta)G(\alpha x_2)\right) \leq e^{-c(\eta)(1-\gamma)mG(\alpha x_2)}$$

and

$$\mathbb{P}\left(\frac{1}{m_1} \sum_{j=1}^{m_1} \mathbb{I}_{\{X_j \leq \alpha x_1\}} \leq (1 - \eta)G(\alpha x_1)\right) \leq e^{-c(\eta)(1-\gamma)mG(\alpha x_1)}.$$

This concludes the proof. \square

As in the case of $\frac{R_m}{m}$ we state a more practical version of Theorem 3 which is aimed to compute the largest ϵ for fixed m, x_1, x_2 . It is in the spirit of Theorem 2.

Theorem 4. Let condition (M) be valid and assume the notation of Theorem 2. Let $m \in \mathbb{N}^+$ be fixed. Let next $x_{\alpha, \gamma} \in (x_1, x_2) \subset (0, 1)$ and x_1 is such that $\frac{\lceil mx_1 \rceil}{m} < x_{\alpha, \gamma}$. Finally, let $\eta \in (0, 1)$ be fixed. Then with

$$\begin{aligned} \epsilon_1 &= \epsilon(x_1, m) = \frac{1}{2} \left(\frac{1}{\alpha} - 1 \right) - \frac{1 - \gamma}{2} H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \\ &= \frac{1 - \gamma}{2} \left(H(\alpha x_{\alpha, \gamma}) - H\left(\alpha \frac{\lceil mx \rceil}{m}\right) \right), \end{aligned} \tag{39}$$

and

$$\begin{aligned} \epsilon_2 &= \epsilon(x_2, m) = \frac{1 - \gamma}{2} H\left(\alpha \frac{\lceil mx \rceil}{m}\right) - \frac{1}{2} \left(\frac{1}{\alpha} - 1 \right) \\ &= \frac{1 - \gamma}{2} \left(H\left(\alpha \frac{\lceil mx \rceil}{m}\right) - H(\alpha x_{\alpha, \gamma}) \right) \end{aligned} \tag{40}$$

we have that

$$\begin{aligned} & \mathbb{P}((1 - \eta)G(\alpha x_1) \leq S_m^* \leq (1 + \eta)G(\alpha x_2)) \\ & \geq \max\{1 - (K_F^M(x_1, m, \gamma, \epsilon_1, \alpha) + K_G^M(x_1, m, 1 - \gamma, \epsilon_1, \alpha)) \\ & \quad - (K_F^M(x_2, m, \gamma, \epsilon_2, \alpha) + K_G^M(x_2, m, 1 - \gamma, \epsilon_2, \alpha)) \\ & \quad - e^{-mc(\eta)(1-\gamma)G(\alpha x_1)} - e^{-mc(\eta)(1-\gamma)G(\alpha x_2)}; 0\} \end{aligned} \tag{41}$$

and $c(\eta) = (1 + \eta) \ln(1 + \eta) - \eta$ is considered in Proposition A1. For a one-sided bound we get the inequality

$$\begin{aligned} & \mathbb{P}((1 - \eta)G(\alpha x_1) \leq S_m^*) \\ & \geq \max\{1 - K_F^M(x_1, m, \gamma, \epsilon_1, \alpha) - K_G^M(x_1, m, 1 - \gamma, \epsilon_1, \alpha) - e^{-mc(\eta)(1-\gamma)G(\alpha x_1)}; 0\}. \end{aligned} \tag{42}$$

Remark 12. Noting Remark 9 we see that the expressions involving K_F^M, K_G^M can be substituted with

$$2 \min \left\{ \sum_{t \in A_\alpha(m, x)} e^{-kT(t)c\left(\frac{\epsilon}{\rho} \frac{t}{T(t)}\right)}, e^{-2\frac{k}{\rho^2}(\epsilon \alpha x)^2} \right\},$$

where T is either F or G .

Proof of Theorem 4. The proof is immediate but we note that the assumption $\frac{\lceil mx_1 \rceil}{m} < x_{\alpha, \gamma}$ and the choice of ϵ_1, ϵ_2 are such that Theorem 2 applies to the right-hand side of (38), which yields an inequality identical to (35). The rest is as in the proof of Theorem 3. \square

3. Workflow

In this section we consider the case when the non-uniform distribution of $(X_j)_{m_0 < j \leq m}$ is Beta with parameters a, b , that is $G(\cdot, a, b)$ is the cdf of $B(a, b)$, where $a < 1$ and $b > 1$ are taken so that condition (M) is satisfied. In this case Lemma 1 is valid and condition (M) holds since elementary calculations yield that $g'_{a,b} < 0$ on $(0, \alpha)$ and $\lim_{x \rightarrow 0} g_{a,b}(x) = \infty$, where $g_{a,b}$ is the probability density function of $B(a, b)$.

For a given number of hypotheses m , level of significance α and proportion of significant hypotheses γ , we are interested in finding a theoretical bound for the probability that the power is between a number l and 1, where of course only $l < G(\alpha x_{\alpha, \gamma}, a, b)$ makes sense to be considered. We shall only provide a step by step description of the implementation and we shall refrain from developing a particular, explicit theoretical example for given values of a, b since the latter involves tedious, non-instructive and lengthy computations. We proceed with describing the workflow that has been implemented numerically in R:

1. we calculate $x_{\alpha, \gamma}$ by solving numerically Equation (22) wherein $G(\cdot) = G(\cdot, a, b)$;
2. using the value of $x_{\alpha, \gamma}$, we calculate the power limit of S_m^* , see (33), i.e., $G(\alpha x_{\alpha, \gamma}, a, b)$, according to (36);
3. for given $l < G(\alpha x_{\alpha, \gamma}, a, b)$, any given η and x_1 is such that $\frac{\lceil mx_1 \rceil}{m} < x_{\alpha, \gamma}$ we have an estimate based on (42), that is

$$\begin{aligned} & \mathbb{P}((1 - \eta)G(\alpha x_1) \leq S_m^*) \\ & \geq \max\{1 - K_F^M(x_1, m, \gamma, \epsilon_1, \alpha) - K_G^M(x_1, m, 1 - \gamma, \epsilon_1, \alpha) - \\ & \quad - e^{-mc(\eta)(1-\gamma)G(\alpha x_1, a, b)}; 0\}; \end{aligned}$$

4. fixing η we attempt to solve (numerically) in x_1 the equation

$$l = (1 - \eta)G(\alpha x_1, a, b);$$

5. if $\frac{\lceil mx_1 \rceil}{m} < x_{\alpha, \gamma}$ then (42) represents a valid theoretical lower bound, that is

$$p(\eta) = \max\{1 - K_F^M(x_1, m, \gamma, \epsilon_1, \alpha) - K_G^M(x_1, m, 1 - \gamma, \epsilon_1, \alpha) - e^{-mc(\eta)(1-\gamma)G(\alpha x_1, a, b)}; 0\}$$

with $\mathbb{P}(l \leq S_m^*) \geq p(\eta)$;

6. finally, we optimize $p(\eta)$ in η and choose η^* so that the estimated probability bound is the largest, that is $\mathbb{P}(l \leq S_m^*) \geq \sup_{\eta \in (0,1)} p(\eta) = p(\eta^*)$.

4. Results

The following Tables 1–4 give empirical results for some values of m, a, b, γ . We have included different cases for the real power limit—taking large, moderate and small values. In all cases the significance level is $\alpha = 0.05$. These tables also show the empirical probability based on numerical simulations with 1 million repetitions. Empirical probability of 1 means that no simulation repetition had power lower than l . Theoretical probability of 0 or empty value of η^* indicate that some of the conditions of the method were not met.

Table 1. Simulation results for B(0.1, 100), number genes = 20,000, proportion significant = 0.1, real power limit = 0.940359.

Lower Bound l	Empirical Probability	Theoretical Estimate $p(\eta^*)$	η^*
0.750	1.000	1.000	0.098
0.760	1.000	1.000	0.097
0.770	1.000	1.000	0.096
0.780	1.000	1.000	0.096
0.790	1.000	1.000	0.095
0.800	1.000	1.000	0.095
0.810	1.000	1.000	0.094
0.820	1.000	1.000	0.094
0.830	1.000	1.000	0.093
0.840	1.000	1.000	0.093
0.850	1.000	0.999	0.085
0.860	1.000	0.996	0.078
0.870	1.000	0.985	0.068
0.880	1.000	0.952	0.058
0.890	1.000	0.873	0.048
0.900	1.000	0.714	0.037
0.910	1.000	0.476	0.027
0.920	1.000	0.204	0.016
0.930	0.967	0.018	0.005
0.940	0.550	0.000	

Larger sample size would result in a smaller parameter a of the Beta distribution, thus squeezing the distribution towards 0 and increasing the statistical power and its theoretical estimates, e.g., as in Tables 2 and 3.

We have carried out numerical simulations for many values of the parameters and based on the results it appears that the theoretical bound works well when l is not close to the real power limit $lG(\alpha x_{\alpha, \gamma}, a, b)$. Link to our R code is provided in the Supplementary Materials section.

Interestingly, when m increases and all other parameters are kept fixed, the expected power decreases for most values of m . This is shown in the following Figure 1. Perhaps this is due to the fact that the set for which we take the maximum in (2) increases its size for most values of m (unless the ceiling function jumps).

We note that as expected when either a is closer to 1 or/and γ is smaller then the theoretical estimates are getting significantly worse. This is natural as the closer the alternative distribution gets to the uniform the harder is to distinguish that the respective random variable X originates from it. The same is valid if the lower bound l is very close to the true limit and it is harder to account for the possible error. All these points reflect natural restrictions when one deals with universal, theoretical bounds, which are seemingly conducted in an optimal way.

Table 2. Simulation results for $B(0.25, 2)$, number genes = 20,000, proportion significant = 0.1, real power limit = 0.233737.

Lower Bound l	Empirical Probability	Theoretical Estimate $p(\eta^*)$	η^*
0.190	1.000	0.779	0.105
0.200	0.997	0.434	0.072
0.210	0.972	0.000	0.001
0.220	0.868	0.000	0.001
0.230	0.626	0.000	0.001

Table 3. Simulation results for $B(0.1, 2)$, number genes = 20,000, proportion significant = 0.1, real power limit = 0.620014.

Lower Bound l	Empirical Probability	Theoretical Estimate $p(\eta^*)$	η^*
0.500	1.000	1.000	0.119
0.510	1.000	1.000	0.117
0.520	1.000	1.000	0.116
0.530	1.000	1.000	0.115
0.540	1.000	0.999	0.106
0.550	1.000	0.996	0.097
0.560	1.000	0.979	0.081
0.570	1.000	0.926	0.067
0.580	1.000	0.786	0.052
0.590	0.994	0.521	0.036
0.600	0.953	0.189	0.02
0.610	0.803	0.000	0.001
0.620	0.511	0.000	

Table 4. Simulation results for $B(0.1, 10)$, number genes = 20,000, proportion significant = 0.1, real power limit = 0.755055.

Lower Bound l	Empirical Probability	Theoretical Estimate $p(\eta^*)$	η^*
0.600	1.000	1.000	0.109
0.610	1.000	1.000	0.108
0.620	1.000	1.000	0.107
0.630	1.000	1.000	0.106
0.640	1.000	1.000	0.105
0.650	1.000	1.000	0.105
0.660	1.000	1.000	0.104
0.670	1.000	0.999	0.096
0.680	1.000	0.995	0.085
0.690	1.000	0.980	0.074
0.700	1.000	0.934	0.062
0.710	1.000	0.821	0.049
0.720	0.999	0.608	0.037
0.730	0.991	0.309	0.024
0.740	0.924	0.048	0.009
0.750	0.693	0.000	0.001

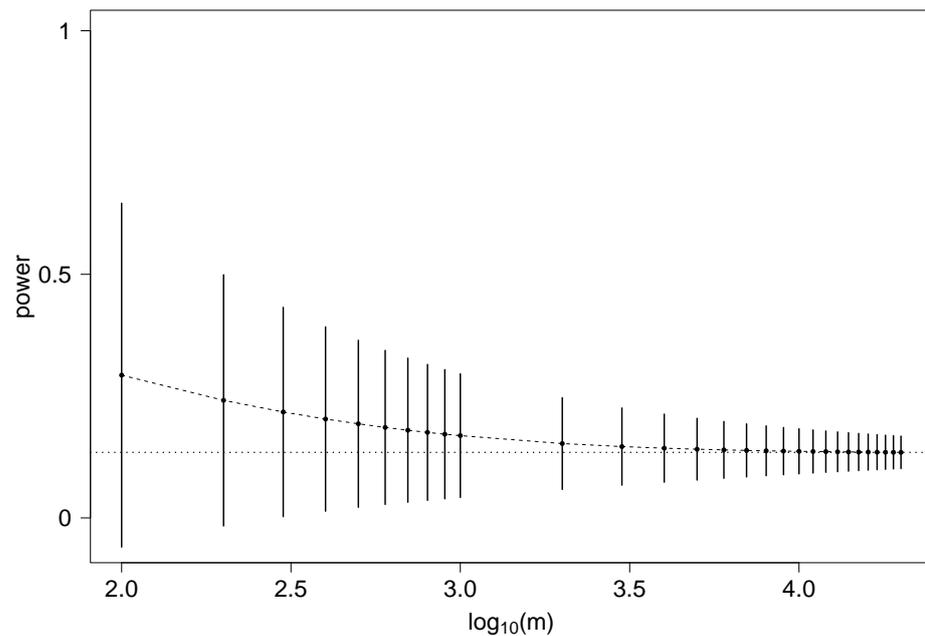


Figure 1. Empirical results when the p -values under the alternative hypothesis have $B(0.5, 100)$ distribution and the proportion of alternative tests is 2%. The x -axis shows the common logarithm of the number of tests m . The y -axis shows the empirical power. The dots depict the mean of the empirical power, the vertical lines are of length one standard deviation below and above the mean. The dotted line shows the asymptotic power.

5. Discussion

5.1. Related Work

There are a number of results concerning the statistical power of the Benjamini–Hochberg procedure that relate to our result. Paper [15] discusses different properties of the procedure and shows that when increasing the number of tests m , while keeping the underlying distributions and the proportion of true significant tests the same, the statistical power under some conditions converges to a limit that can be found as a solution of a particular equation, see (3.8) of Corollary 3.3 of that article. A step forward is made in [13] wherein it is shown that the proportions related to the power of the test obey even a law of the iterated logarithm. Other notable contributions amongst others include the papers [18] which deal with various properties of the Benjamini–Hochberg procedure, including a Donsker-type of convergence. They are discussed in more detail in Remark 5. We wish to highlight explicitly that to our knowledge there are no bounds, albeit conservative, for the probability to uncover a given proportion of the truly significant tests (that is S_n/n) with respect to the limit of S_n/n , as n goes to infinity. We achieve this by evaluating the speed of convergence, whereas the results under more general assumptions, including under weak dependence between the tests, consider limit results. This does not allow for the derivation of any concrete intervals, depending solely on the set of parameters, which give a certain probability for the proportion of correctly identified significant tests.

While we argue below that for genetic studies the assumption of independence is not binding at all, we wish to mention some achievements under the weak dependence assumption since it is natural to pursue as a future development the extension of our results to this scenario. The limit for the power of the Benjamini–Hochberg procedure is proved in [15]. An interesting contribution is [12] which shows that for a large number of tests the FDR is controlled even when there is some degree of dependence between the test statistics. A conservative estimator for the proportion of false nulls among the hypotheses (the quantity $1 - \gamma$ in our work) is found (the proof there uses the Dvoretzky–Kiefer–Wolfowitz inequality [16], as we do for different purposes in our work). Then the estimator is used in

a plug-in procedure in order to boost the power of the procedure while keeping the FDR controlled. An estimator for γ may precede the application of our procedure.

We are not interested in improving the procedure itself, but rather than that in estimating the power of the original procedure, that as we mentioned in the Introduction section, is the most widely used among the similar procedures.

As mentioned in [20], the results in [12] assume that R_n/n is asymptotically larger than zero, where R_n is the number of rejected tests out of all n tests (the R notation is consistent with ours through this work). A counterexample is given of a case in which R_n/n converges asymptotically to zero with positive probability and in which the FDR is larger than α . Such situation does not appear to be relevant to the RNA-seq data used for single-step differential expression methods that prompted our work. Moreover, paper [12] is concerned with controlling the FDR rather than estimates of the power, which is the scope of our work in which we are interested in the power of the procedure that the scientists predominantly use, regardless of whether the FDR is controlled.

5.2. Considerations about Our Work

Because this is the first result of this exact type that we are aware of, we start our considerations under independence assumption. We want to point out that, as mentioned above, there are more general results under convergence, but they do not include explicit bounds for given set of parameters. The independence assumption holds approximately in the case of RNA-seq expression level analysis, where there are groups of genes that are correlated within the group, but most pairs of genes are weakly or not correlated. Even in the case of weak dependence, having real-world RNA-seq expression data, one cannot reliably estimate or replicate the dependence unless the sample size is much larger than the one of the currently available datasets. In addition, more complicated assumptions for the structure of dependence would make the numerical procedure impossible to set up as there would be too many parameters in the input and these parameters cannot be easily estimated for small sample sizes that are typical for RNA-seq experiments.

Theoretical results under more complicated assumptions for the dependence could be a subject of a follow-up studies. Numerical simulations of RNA-seq data under different basic types of weak dependencies could also be considered in subsequent studies.

It is important to also emphasize that our considerations differ from the basic situations in which the statistical power is studied. In a case of a simple test the statistical power is typically considered as a function of the sample size. However, in a case of a complex tests such as the ones cited above, often it is not possible to estimate the distribution of the test statistic, even for a single test in the multiple comparison setup. In order to be able to model a general setup we model the test statistic under the alternative hypothesis using the Beta family of distributions. In order to use our current results, one would first have to approximately or exactly estimate the parameters of the Beta distribution as function of the sample size of the particular test. Because such estimates are test-specific and cannot be studied in general, this would be a subject of another article. One would also have to estimate the proportion of significant tests and, having achieved that, one would apply our results for that proportion and those parameters of the Beta distribution, taking into account the possible values of the number of tests. For example, in the differential expression analysis methods, mentioned above, the sample size corresponds to the number of individuals and the number of tests may correspond to the number of genes.

Supplementary Materials: The R script and instructions are available at the following page: <http://www.math.bas.bg/~palejev/BH>.

Funding: D.P. was supported by the European Union Horizon 2020 WIDESPREAD-2018-2020 TEAMING Phase 2 programme under Grant Agreement No. 857155 and Operational Programme Science and Education for Smart Growth under Grant Agreement No. BG05M2OP001-1.003-0002-C01. M.S. was partially supported by the financial funds allocated to the Sofia University “St. Kliment Ohridski”, grant No. 80-10-87/2021. The computational resources and infrastructure were provided

by NCHDC—part of the Bulgarian National Roadmap of RIs, with the financial support by Grant No DOI—387/18.12.2020.

Acknowledgments: The extensive numerical simulations were done on the Avitohol supercomputer that is described in [21].

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- FDR False Discovery Rate
- PET Positron Emission Tomography
- CT Computerized Tomography
- fMRI Functional Magnetic Resonance Imaging

Appendix A. Petrov’s Inequality

In this part we state and outline the proof of a celebrated inequality which can be found in the monograph [22].

Proposition A1. *Let X is binomially distributed with parameters n, p . Then we have that for any $\epsilon > 0$*

$$\mathbb{P}(X - np > \epsilon np) \leq e^{-npc(\epsilon)}, \quad \mathbb{P}(np - X > \epsilon np) \leq e^{-npc(\epsilon)} \tag{A1}$$

where $c(\epsilon) = (1 + \epsilon) \ln(1 + \epsilon) - \epsilon, c : (0, \infty) \mapsto (0, \infty)$ is an increasing function and the limit $\lim_{\epsilon \rightarrow 0} 2c(\epsilon)/\epsilon^2 = 1$ holds. Finally, $c(\epsilon) \leq \epsilon^2$ and $c(\epsilon) \geq \min\left(1, \frac{\epsilon^2}{2e}\right)$ for any $\epsilon > 0$.

Proof. We consider

$$\mathbb{P}(X > (1 + \epsilon)np) \leq e^{-\lambda np(1+\epsilon)} \mathbb{E}\left[e^{\lambda X}\right] = e^{n \ln(1-p+pe^\lambda) - \lambda np(1+\epsilon)},$$

where the latter is simply the Markov’s inequality with some $\lambda > 0$. Then using $\ln(1 + \epsilon) \leq \epsilon, \epsilon > 0$, we deduct that

$$\mathbb{P}(X > (1 + \epsilon)np) \leq e^{np(e^\lambda - 1 - \lambda(1+\epsilon))}.$$

The minimum of $f(\lambda) = e^\lambda - 1 - \lambda(1 + \epsilon)$ is attained at $\lambda = \ln(1 + \epsilon)$ and thus we conclude that

$$\mathbb{P}(X > (1 + \epsilon)np) \leq e^{-np((1+\epsilon) \ln(1+\epsilon) - \epsilon)}.$$

Similar computation works for $\mathbb{P}(X > (1 - \epsilon)np)$. Clearly, setting the function

$$c(\epsilon) = (\epsilon - (1 + \epsilon) \times \ln(1 + \epsilon))$$

we arrive at the limit

$$\lim_{\epsilon \rightarrow 0} 2 \frac{(\epsilon - (1 + \epsilon) \ln(1 + \epsilon))}{\epsilon^2} = 1$$

and $(1 + \epsilon) \ln(1 + \epsilon) - \epsilon$ is increasing in ϵ . Using again $\ln(1 + \epsilon) \leq \epsilon, \epsilon > 0$, we also conclude that $c(\epsilon) \leq \epsilon^2$ for any $\epsilon > 0$. Considering the properties of the function $f(\epsilon) = c(\epsilon) - \frac{\epsilon^2}{2e}$ for $\epsilon \in (0, e - 1)$ and noting that $c(\epsilon) \geq 1$ for $\epsilon > e - 1$ we deduct that $c(\epsilon) \geq \min\left(1, \frac{\epsilon^2}{2e}\right)$. \square

References

1. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 289–300. [[CrossRef](#)]
2. Love, M.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)]
3. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [[CrossRef](#)] [[PubMed](#)]
4. Trapnell, C.; Hendrickson, D.G.; Sauvageau, M.; Goff, L.; Rinn, J.L.; Pachter, L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* **2012**, *31*, 46–53. [[CrossRef](#)]
5. Chalkidou, A.; O'Doherty, M.J.; Marsden, P.K. False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review. *PLoS ONE* **2015**, *10*, e0124165. [[CrossRef](#)]
6. Bennett, C.M.; Wolford, G.L.; Miller, M.B. The principled control of false positives in neuroimaging. *Soc. Cogn. Affect. Neurosci.* **2009**, *4*, 417–422. [[CrossRef](#)]
7. Miller, C.J.; Genovese, C.; Nichol, R.C.; Wasserman, L.; Connolly, A.; Reichart, D.; Hopkins, A.; Schneider, J.; Moore, A. Controlling the False-Discovery Rate in Astrophysical Data Analysis. *Astron. J.* **2001**, *122*, 3492. [[CrossRef](#)]
8. Benjamini, Y.; Yekutieli, D. The Control of the False Discovery Rate in Multiple Testing under Dependency. *Ann. Stat.* **2001**, *29*, 1165–1188. [[CrossRef](#)]
9. Storey, J.D.; Taylor, J.E.; Siegmund, D. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: A unified approach. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2004**, *66*, 187–205. [[CrossRef](#)]
10. Pollard, K.S.; van der Laan, M.J. Choice of a null distribution in resampling-based multiple testing. *J. Stat. Plan. Inference* **2004**, *125*, 85–100. [[CrossRef](#)]
11. Van Noorden, R.; Maher, B.; Nuzzo, R. The top 100 papers. *Nature* **2014**, *514*, 550–553. [[CrossRef](#)]
12. Farcomeni, A. Some Results on the Control of the False Discovery Rate under Dependence. *Scand. J. Stat.* **2007**, *34*, 275–297. [[CrossRef](#)]
13. Chi, Z. On the performance of FDR control: Constraints and a partial solution. *Ann. Statist.* **2007**, *35*, 1409–1431. [[CrossRef](#)]
14. Ferreira, J.A.; Zwiderman, A.H. Approximate Power and Sample Size Calculations with the Benjamini-Hochberg Method. *Int. J. Biostat* **2006**, *2*, 8. [[CrossRef](#)]
15. Ferreira, J.A.; Zwiderman, A.H. On the Benjamini—Hochberg method. *Ann. Statist.* **2006**, *34*, 1827–1849. [[CrossRef](#)]
16. Dvoretzky, A.; Kiefer, J.; Wolfowitz, J. Asymptotic Minimax Character of the Sample Distribution Function and of the Classical Multinomial Estimator. *Ann. Math. Stat.* **1956**, *27*, 642–669. [[CrossRef](#)]
17. Pounds, S.; Morris, S.W. Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values. *Bioinformatics* **2003**, *19*, 1236–1242. [[CrossRef](#)] [[PubMed](#)]
18. Neuvial, P. Asymptotic properties of false discovery rate controlling procedures under independence. *Electron. J. Stat.* **2008**, *2*, 1065–1110. [[CrossRef](#)]
19. Finner, H.; Dickhaus, T.; Roters, M. On the false discovery rate and an asymptotically optimal rejection curve. *Ann. Statist.* **2009**, *37*, 596–618. [[CrossRef](#)]
20. Gontscharuk, V.; Finner, H. Asymptotic FDR control under weak dependence: A counterexample. *Stat. Probab. Lett.* **2013**, *83*, 1888–1893. [[CrossRef](#)]
21. Atanassov, E.; Gurov, T.; Karaivanova, A.; Ivanovska, S.; Durchova, M.; Dimitrov, D. On the parallelization approaches for Intel MIC architecture. *AIP Conf. Proc.* **2016**, *1773*, 070001. [[CrossRef](#)]
22. Petrov, V.V. *Sums of Independent Random Variables*; Brown, A.A., Translator; Ergebnisse der Mathematik und ihrer Grenzgebiete; Springer: Berlin/Heidelberg, Germany, 1975; Volume 82.