Various New Inequalities for Beta Distributions*

Alexander Henzi and Lutz Dümbgen
ETH Zürich and University of Bern

December 2022

Abstract

This note provides some new inequalities and approximations for beta distributions, including tail inequalities, exponential inequalities of Hoeffding and Bernstein type, Gaussian inequalities and approximations.

Keywords: Exponential inequality, gamma distribution, Gaussian approximation, tail inequality, density ratio.

MSC2020 subject classifications: Primary 62E17; 60E05; 60E15, Secondary 33B20.

1 Introduction

Beta distributions play an important role in statistics and probability theory (Gupta and Nadarajah, 2004), and they occur in various scientific fields (Skorski, 2021). A frequent obstacle in problems involving beta distributions is the lack of analytic expressions for their distribution function, the normalized incomplete beta function. Therefore one often resorts to inequalities and approximations, as, for example, in the proofs of Dimitriadis et al. (2022, Theorem 4.1) and Dümbgen and Wellner (2022, Lemma S.8).

This paper provides some new inequalities for the beta distribution $\mathrm{Beta}(a, b)$ with parameters $a, b > 0$, its distribution function $B_{a,b}$, survival function $\bar{B}_{a,b} = 1 - B_{a,b}$ and density function $\beta_{a,b}$ on $[0, 1]$. The latter is given by

$$\beta_{a,b}(x) \;:=\; B(a,b)^{-1} x^{a-1}(1-x)^{b-1}, \quad x \in (0,1),$$

wheere $B(a,b) \;:=\; \int_0^1 x^{a-1}(1-x)^{b-1}\,\mathrm{d}x \;=\; \Gamma(a)\Gamma(b)/\Gamma(a+b)$, and $\Gamma(\cdot)$ denotes the gamma function. In Section 2, we refine the lower and upper bounds for $B_{a,b}$ and $\bar{B}_{a,b}$ by Segura (2016) which are particularly accurate in the tails of $\mathrm{Beta}(a, b)$. As a by-product we obtain refinements of Segura's (2014) bounds for the gamma distribution and survival functions. In Section 3 we present new exponential inequalities which are stronger than previously known inequalities of

*This is an extended version of the paper "Some New Inequalities for Beta Distributions", published in *Statistics & Probability Letters* ***195*** *(2023)*

source: https://doi.org/10.48350/177383 | downloaded: 17.2.2023

Dümbgen (1998), Marchal and Arbel (2017) and Skorski (2021). Section 4 presents inequalities for $B_{a,b}$ and $\bar{B}_{a,b}$ in terms of Gaussian distribution functions. Finally, Section 5 discusses the approximation of the symmetric distribution $\beta_{a,a}$ by Gaussian densities with mean $1/2$ in the spirit of Dümbgen et al. (2021). Most proofs are deferred to Section 6.

2 Sharp tail inequalities

In what follows, let $p := a/(a+b)$, the mean of $\mathrm{Beta}(a,b)$. In a general setting including noncentral beta distributions, Segura (2016, inequalities (27), (29), (30)) uses extensions of l'Hopital's rule to derive inequalities for $B_{a,b}$ and $\bar{B}_{a,b}$. For symmetry reasons, we only consider $B_{a,b}$, because $\bar{B}_{a,b}(x) = B_{b,a}(1-x)$. We rephrase Segura's inequalities in terms of the ratio

$$Q_{a,b}(x) \ := \ \frac{B_{a,b}(x)}{x^a/[aB(a,b)]}.$$

This is motivated by the fact that $\beta_{a,b}(x) = B(a,b)^{-1}x^{a-1}(1+O(x))$ and thus $B_{a,b}(x) = x^a/[aB(a,b)](1+O(x))$ as $x \to 0$. The goal is to find upper and lower bounds for $Q_{a,b}(x)$ approaching 1 as $x \to 0$. Now, for $x \in (0,1)$,

$$(1) \qquad (1-x)^b(1+c_{a,b}x) \ \le \ Q_{a,b}(x) \ \le \ \frac{(1-x)^b}{(1-x/p)^+},$$

where $c_{a,b} := (a+b)/(a+1)$. Numerical examples reveal that these inequalities are rather accurate unless x is close to or larger than p. Our first contribution is an improvement of Segura's bounds. In particular, the upper bound remains valid if x/p is replaced with the strictly smaller term $\max\{c_{a,b},1\}x$. The results are stated in terms of the following auxiliary functions:

$$\begin{aligned}
q_{a,b}^{(1)}(x) \ &:= \ \Big(1 - \frac{ax}{a+1}\Big)^{b-1}, \\
q_{a,b}^{(2)}(x) \ &:= \ \frac{a(1-x)^{b-1}+1}{a+1} - \frac{a(b-1)(b-2)x^2(1-x)^{(b-3)^+}}{2(a+1)(a+2)}, \\
q_{a,b}^{(3)}(x) \ &:= \ \frac{(1-x)^b}{(1-c_{a,b}x)^+}.
\end{aligned}$$

Theorem 1. *For $x \in (0,1)$,*

$$(1-x)^b(1+c_{a,b}x) \ < \ Q_{a,b}^L(x) \ \le \ Q_{a,b}(x) \ \le \ Q_{a,b}^U(x) \ \le \ \frac{(1-x)^b}{(1-\max\{c_{a,b},1\}x)^+},$$

where

$$Q_{a,b}^L(x) \ := \ \begin{cases} q_{a,b}^{(1)}(x) & \text{if } b \notin (1,2), \\ q_{a,b}^{(2)}(x) & \text{if } b \in [1,2], \end{cases}$$

$$Q_{a,b}^U(x) \ := \ \begin{cases} q_{a,b}^{(2)}(x) & \text{if } b \le 1, \\ q_{a,b}^{(1)}(x) & \text{if } b \in [1,2], \\ \min\{q_{a,b}^{(2)}(x), q_{a,b}^{(3)}(x)\} & \text{if } b > 2. \end{cases}$$

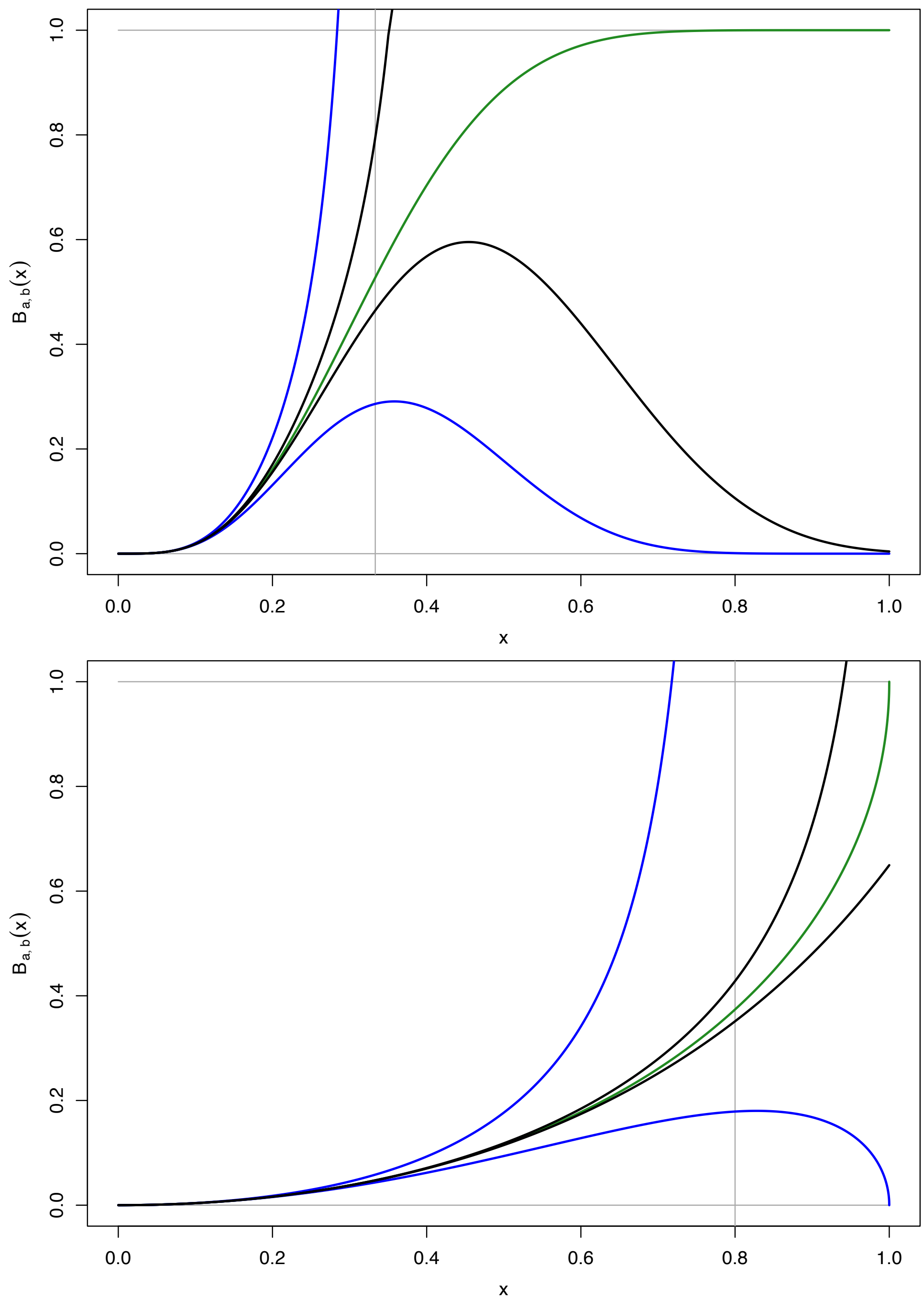

Figure 1: Inequalities for $B_{a,b}$ when $(a, b) = (4, 8)$ (upper panel) and $(a, b) = (2, 0.5)$ (lower panel). The green line shows $B_{a,b}$, the blue lines are Segura's (2016) bounds resulting from (1), and the black lines are the bounds via Theorem 1. The vertical line indicates the mean p.

Figure 1 illustrates the bounds for $B_{a,b}$ resulting from (1) and Theorem 1 in case of $(a,b) = (4,8),(2,0.5)$.

Remark 2. The new inequalites for $Q_{a,b}$ are equalities in the case of $b \in \{1,2\}$, because $q^{(1)}_{a,1}(x) = q^{(2)}_{a,1}(x) = Q_{a,1}(x) = 1$ and $q^{(1)}_{a,2}(x) = q^{(2)}_{a,2}(x) = Q_{a,2}(x) = 1 - ax/(a+1)$. Moreover, the upper bound is exact for $b = 3$, because $q^{(2)}_{a,3}(x) = Q_{a,3}(x) = 1 - 2ax/(a+1) + ax^2/(a+2)$.

Remark 3. Note that the ratio $Q_{a,b}$ as well as the bounds $Q^L_{a,b}, Q^U_{a,b}$ are equal to $1 - d_{a,b}x + O(x^2)$ as $x \to 0$, where $d_{a,b} = (b-1)a/(a+1)$. The lower bound in (1) has the same property, but the upper bound does not.

Gamma distributions. There is a rich literature about inequalities for gamma distribution and survival functions, see, for instance, Qi and Mei (1999), Neuman (2013), Segura (2014) and Pinelis (2020). We just illustrate that our bounds in Theorem 1 yield a connection to that literature. It is well-known that for a random variable $X_{a,b} \sim \mathrm{Beta}(a,b)$, the rescaled variable $bX_{a,b}$ converges in distribution to a gamma random variable with shape parameter a and scale parameter 1 as $b \to \infty$. Denoting the corresponding distribution and survival function with G_a and $\bar{G}_a = 1 - G_a$, respectively, we have $G_a(x) = \lim_{b\to\infty} B_{a,b}(x/b)$, and one can deduce from Theorem 1 the following bounds.

Corollary 4. *For* $a, x > 0$,

$$\frac{x^a e^{-ax/(a+1)}}{a\Gamma(a)} \;\le\; G_a(x) \;\le\; \frac{x^a}{a\Gamma(a)} \cdot \begin{cases} \dfrac{ae^{-x}+1}{a+1} - \dfrac{ax^2 e^{-x}}{2(a+1)(a+2)}, \\[2mm] \dfrac{e^{-x}}{(1 - x/(a+1))^+}, \end{cases}$$

$$\frac{(x + 1_{[a \notin (1,2)]})^{a-1} e^{-x}}{\Gamma(a)} \le \bar{G}_a(x) \;\le\; \begin{cases} \dfrac{(x + 1_{[a>1]})^{a-1} e^{-x}}{\Gamma(a)} & \text{if } a \le 2, \\[2mm] \dfrac{x^a e^{-x}}{\Gamma(a)(x - a + 1)^+} & \text{if } a > 2, \\[2mm] e^{-x}(x^2/2 + x + 1) & \text{if } a = 3. \end{cases}$$

The lower bound for $G_a(x)$ is already known from Neuman (2013, Theorem 4.1), and the upper bounds for $G_a(x)$ are a combination of Segura (2014, Theorem 10, part 3) and a slight improvement of Neuman (2013, Theorem 4.1). The lower bounds for $\bar{G}_a(x)$ are equalities if $a \in \{1,2\}$, and the upper bounds if $a \in \{1,2,3\}$. Our lower bound for $\bar{G}_a(x)$ extends the lower bound of Segura (2014, Theorem 10, part 4) to $a < 1$, and it is stronger than the latter for $a > 2$. Our upper bound for $\bar{G}_a(x)$ extends the upper bound of Segura (2014, Theorem 10, part 6) to $a < 1$, and it is stronger than the latter if $1 < a \le 2$.

3 Exponential inequalities

Although the upper bounds in Theorem 1 are numerically rather accurate in the tails, they can diverge to ∞ at $x = p$ as $a, b \to \infty$. Moreover, it is sometimes desirable to have bounds for $\log B_{a,b}(x)$ and $\log \bar{B}_{a,b}(x)$ in terms of simple, maybe rational, functions of x. Numerous exponential tail inequalities for $B_{a,b}$ and $\bar{B}_{a,b}$ have been derived already. We start with one particular result of Dümbgen (1998, Proposition 2.1). For $x \in [0, 1]$ let

$$K(p, x) \ := \ p \log\Bigl(\frac{p}{x}\Bigr) + (1-p)\log\Bigl(\frac{1-p}{1-x}\Bigr) \in [0, \infty].$$

This function $K(p, \cdot)$ is strictly convex with minimum $K(p, p) = 0$. For arbitrary $x \in [0, 1]$,

$$\text{(2)} \qquad \frac{x^a(1-x)^b}{p^a(1-p)^b} = \exp\bigl(-(a+b)K(p,x)\bigr) \ \geq \ \begin{cases} B_{a,b}(x) & \text{if } x \leq p, \\ \bar{B}_{a,b}(x) & \text{if } x \geq p. \end{cases}$$

In case of $a \geq 1$ or $b \geq 1$, these inequalities can be improved as follows.

Theorem 5. *Suppose that $a \geq 1$. Then for $p_r := (a-1)/(a+b-1) < p$ and $x \in [p_r, 1]$,*

$$\bar{B}_{a,b}(x) \begin{cases} \leq \ \dfrac{x^{a-1}(1-x)^b}{p_r^{a-1}(1-p_r)^b} \ = \ \exp\bigl(-(a+b-1)K(p_r, x)\bigr), \\ \geq \ \dfrac{x^{a-1}(1-x)^b}{bB(a,b)}. \end{cases}$$

Suppose that $b \geq 1$. Then for $p_\ell := a/(a+b-1) > p$ and $x \in [0, p_\ell]$,

$$B_{a,b}(x) \begin{cases} \leq \ \dfrac{x^{a}(1-x)^{b-1}}{p_\ell^{a}(1-p_\ell)^{b-1}} \ = \ \exp\bigl(-(a+b-1)K(p_\ell, x)\bigr), \\ \geq \ \dfrac{x^{a}(1-x)^{b-1}}{aB(a,b)}. \end{cases}$$

Remark 6. At first glance, the upper bounds in Theorem 5 seem to be weaker than the ones in (2), at least in the tail regions, because the factor $a + b - 1$ is strictly smaller than $a + b$. But elementary algebra reveals that in case of $a \geq 1$,

$$\begin{aligned} &(a+b-1)K(p_r, x) - (a+b)K(p,x) \\ &\quad = \ \log\Bigl(\frac{x}{p}\Bigr) + (a+b-1)\log\Bigl(1 + \frac{1}{a+b-1}\Bigr) - (a-1)\log\Bigl(1 + \frac{1}{a-1}\Bigr) \\ &\quad > \ 0 \quad \text{for } x \in [p, 1), \end{aligned}$$

because $h(y) := y\log(1 + 1/y)$ (with $h(0) := 0$) is strictly increasing in $y \geq 0$. Analogously, if $b \geq 1$, then

$$\begin{aligned} &(a+b-1)K(p_\ell, x) - (a+b)K(p,x) \\ &\quad = \ \log\Bigl(\frac{1-x}{1-p}\Bigr) + (a+b-1)\log\Bigl(1 + \frac{1}{a+b-1}\Bigr) - (b-1)\log\Bigl(1 + \frac{1}{b-1}\Bigr) \\ &\quad > \ 0 \quad \text{for } x \in (0, p]. \end{aligned}$$

Thus the bounds in Theorem 5 are strictly smaller than the bounds in (2). This is illustrated in Figure 2 for $(a, b) = (4, 8)$.

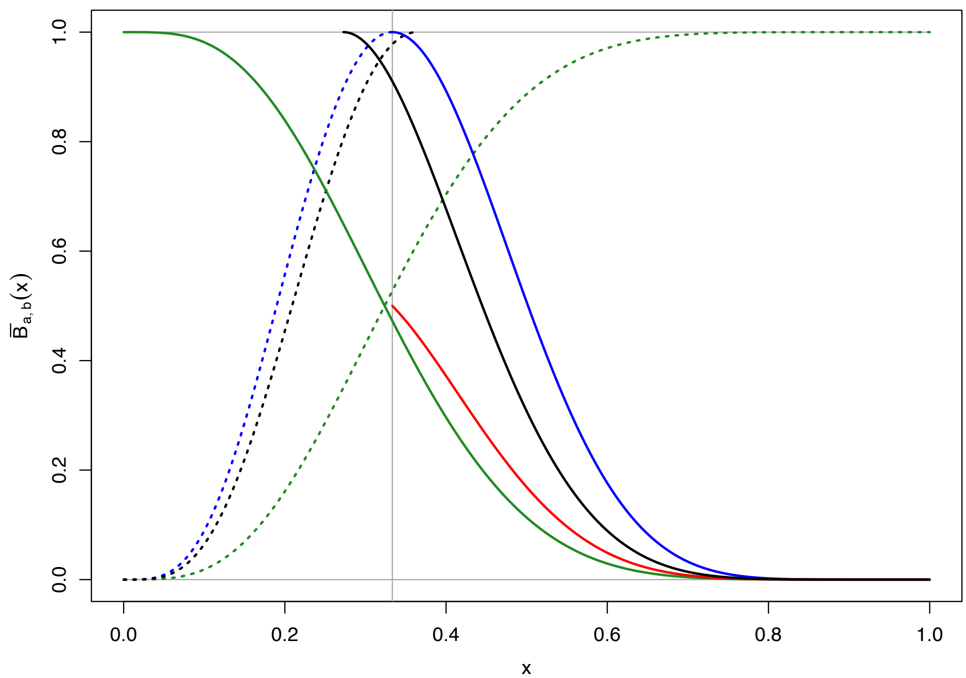

Figure 2: Exponential tail inequalites for $\mathrm{Beta}(a, b)$ when $(a, b) = (4, 8)$. The green line shows $\bar{B}_{a,b}$, the black line is its upper bound from Theorem 5, and the blue line is its upper bound from (2). One also sees the distribution function $B_{a,b}$ and its bounds as dotted lines. The additional red line is the upper bound (3) from Remark 7.

Remark 7. The upper bound for $\bar{B}_{a,b}$ in Theorem 5 can be improved substantially if $1 \le a \le b$. Indeed, the proof of Theorem 5 shows that for arbitrary $0 < x_o \le x \le 1$,

$$\bar{B}_{a,b}(x) \ \le \ \frac{\bar{B}_{a,b}(x_o)}{x_o^{a-1}(1-x_o)^b} x^{a-1}(1-x)^b.$$

Specifically, it is well-known that $\mathrm{Median}(\mathrm{Beta}(a, b)) \le p$, see Groeneveld and Meeden (1977), so $\bar{B}_{a,b}(p) \le 1/2$ and for $x \in [p, 1]$,

(3) $$\bar{B}_{a,b}(x) \ \le \ \frac{x^{a-1}(1-x)^b}{2p^{a-1}(1-p)^b}.$$

The latter bound is strictly smaller than the upper bound of Theorem 5 (restricted to $x \in [p, 1]$), provided that $2p^{a-1}(1-p)^b > p_r^{a-1}(1-p_r)^b$, and this is equivalent to $h(a-1) > h(a+b-1) - \log(2)$ with the increasing function $h(y) = y\log(1 + 1/y)$, $y > 0$. Since $h(a+b-1) < \lim_{y\to\infty} h(y) = 1$, a sufficient condition is that $h(a-1) \ge 1 - \log(2)$, which is fulfilled for $a \ge 1.152$.

The inequalities in Theorem 5 imply Bernstein and Hoeffding type exponential inequalities. It follows from Dümbgen and Wellner (2022, Lemma S.12) and the well-known inequality $z(1-z) \le 1/4$ for $z \in \mathbb{R}$, that

(4) $$K(q, x) \ \ge \ \frac{(x-q)^2}{2(2x/3 + q/3)(1 - 2x/3 - q/3)} \ \ge \ 2(x-q)^2$$

for $q, x \in [0,1]$, where $K(0,x) := -\log(1-x)$ and $K(1,x) := -\log(x)$. This leads to the following inequalities:

Corollary 8. *If $a \ge 1$, then for $x \in [p_r, 1]$,*

$$\begin{aligned}\bar{B}_{a,b}(x) \ &\le\ \exp\Big(-\frac{(a+b-1)(x-p_r)^2}{2(2x/3+p_r/3)(1-2x/3-p_r/3)}\Big)\\ &\le\ \exp\big(-2(a+b-1)(x-p_r)^2\big).\end{aligned}$$

If $b \ge 1$, then for $x \in [0, p_\ell]$,

$$\begin{aligned}B_{a,b}(x) \ &\le\ \exp\Big(-(a+b-1)\frac{(x-p_\ell)^2}{2(2x/3+p_\ell/3)(1-2x/3-p_\ell/3)}\Big)\\ &\le\ \exp\big(-2(a+b-1)(x-p_\ell)^2\big).\end{aligned}$$

Further tail and concentration inequalities for the Beta distribution have been derived by Marchal and Arbel (2017) and Skorski (2021). Marchal and Arbel (2017) prove that $\mathrm{Beta}(a,b)$ is subgaussian with a variance parameter that is the solution of an equation involving hypergeometric functions. An analytic upper bound for the variance parameter is $(4(a+b+1))^{-1}$, which implies the tail inequalities

$$\exp\big(-2(a+b+1)(x-p)^2\big) \ \ge\ \begin{cases} B_{a,b}(x) & \text{if } x \le p,\\ \bar{B}_{a,b}(x) & \text{if } x \ge p.\end{cases}$$

These bounds are weaker than the one-sided bounds in Corollary 8. For the right tails, the difference

$$(a+b-1)(x-p_r)^2 - (a+b+1)(x-p)^2$$

is strictly concave in x with value $b^2/[(a+b)^2(a+b-1)] > 0$ for $x \in \{p, 1\}$. Analogously, for the left tails, the difference

$$(a+b-1)(x-p_\ell)^2 - (a+b+1)(x-p)^2$$

is strictly concave in x with value $a^2/[(a+b)^2(a+b-1)] > 0$ for $x \in \{0, p\}$. Skorski (2021) derives a Bernstein type inequality. With the parameters

$$v^2 \ :=\ \frac{p(1-p)}{a+b+1}, \quad c \ :=\ \max\Big(\frac{|1-2p|}{a+b+2}, \sqrt{\frac{p(1-p)}{a+b+2}}\Big),$$

he shows that for $X \sim \mathrm{Beta}(a,b)$ and $\varepsilon \ge 0$,

$$P(\pm(X-p) \ge \varepsilon) \ \le\ \exp\Big(-\frac{\varepsilon^2}{2(v^2+c\varepsilon)}\Big).$$

The next result shows that our bounds imply a stronger version of these inequalities if $a, b \ge 1$.

Corollary 9. *Let $a, b \ge 1$. Then for $x \in [p, 1]$,*

$$\bar{B}_{a,b}(x) \ \le\ \exp\Big(-\frac{(a+b+1)(x-p)^2}{2p(1-p)+(4/3)(1-2p)(x-p)}\Big),$$

and for $x \in [0, p]$,

$$B_{a,b}(x) \ \le \ \exp\Bigl(-\frac{(a+b+1)(x-p)^2}{2p(1-p)+(4/3)(2p-1)(p-x)}\Bigr).$$

With the notation of Skorski (2021), our upper bounds read

$$P(\pm(X-p) \ge \varepsilon) \ \le \ \exp\Bigl(-\frac{\varepsilon^2}{2(v^2 \pm \tilde{c}\varepsilon)}\Bigr)$$

with v^2 as before and $\tilde{c} = (2/3)(1-2p)/(a+b+1)$. In particular,

$$|\tilde{c}| \ = \ \frac{2(a+b+2)}{3(a+b+1)}\frac{|1-2p|}{a+b+2} \ \le \ \frac{2(a+b+2)}{3(a+b+1)}c.$$

Since $a, b \ge 1$, the factor $2(a+b+2)/[3(a+b+1)]$ is at most $8/9$ and converges to $2/3$ as $a+b \to \infty$.

4 Gaussian tail inequalities

Now suppose that $a, b > 1$. With $p_o := (a-1)/(a+b-2) \in (0,1)$, the density $\beta_{a,b}$ may be written as

$$\log \beta_{a,b}(x) \ = \ \log \beta_{a,b}(p_o) - (a+b-2)K(p_o, x),$$

whereas the probability density $\phi_{p_o,\sigma}$ of $\mathcal{N}(p_o, \sigma^2)$ with $\sigma := (4(a+b-2))^{-1/2}$ satisfies

$$\log \phi_{p_o,\sigma}(x) \ = \ \log \phi_{p_o,\sigma}(p_o) - 2(a+b-2)(x-p_o)^2.$$

In particular, $\rho := \log(\beta_{a,b}/\phi_{p_o,\sigma})$ satisfies

$$\rho'(x) \ = \ (a+b-2)(x-p_o)\bigl(4 - 1/[x(1-x)]\bigr),$$

and since $x(1-x) \le 1/4$, $\rho(x)$ is monotone decreasing in $x \ge p_o$ and monotone increasing in $x \le p_o$, where $\beta_{a,b} := 0$ on $\mathbb{R} \setminus (0,1)$. Consequently, for $x \ge p_o$,

$$\begin{aligned}
\bar{B}_{a,b}(x) \ &\le \ \frac{\bar{B}_{a,b}(x)}{\bar{B}_{a,b}(p_o)} \\
&= \ \frac{\int_x^\infty e^{\rho(t)}\phi_{p_o,\sigma}(t)\,\mathrm{d}t}{\int_{p_o}^x e^{\rho(t)}\phi_{p_o,\sigma}(t)\,\mathrm{d}t + \int_x^\infty e^{\rho(t)}\phi_{p_o,\sigma}(t)\,\mathrm{d}t} \\
&\le \ \frac{e^{\rho(x)}\int_x^\infty \phi_{p_o,\sigma}(t)\,\mathrm{d}t}{e^{\rho(x)}\int_{p_o}^x \phi_{p_o,\sigma}(t)\,\mathrm{d}t + e^{\rho(x)}\int_x^\infty \phi_{p_o,\sigma}(t)\,\mathrm{d}t} \\
&= \ \frac{\mathcal{N}(p_o,\sigma^2)([x,\infty))}{\mathcal{N}(p_o,\sigma^2)([p_o,\infty))} \\
&= \ 2\Phi\bigl(-2\sqrt{a+b-2}(x-p_o)\bigr).
\end{aligned}$$

Analogous arguments apply to $B_{a,b}(x)$ for $x \le p_o$, and we obtain the following bounds.

Lemma 10. *For* $a, b > 1$ *and* $p_o = (a-1)/(a+b-2)$,

$$\begin{aligned}
\bar{B}_{a,b}(x) \ &\le \ 2\Phi\bigl(-2\sqrt{a+b-2}(x-p_o)\bigr) \quad \textit{for } x \ge p_o, \\
B_{a,b}(x) \ &\le \ 2\Phi\bigl(2\sqrt{a+b-2}(x-p_o)\bigr) \quad \textit{for } x \le p_o.
\end{aligned}$$

5 Gaussian approximation of $\mathrm{Beta}(a,a)$

Inspired by Dümbgen et al. (2021), we want to compare the densities $\beta_{a,a}$ with the density $\phi_{1/2,\sigma}$ of $\mathcal{N}(1/2,\sigma^2)$ for various choices of $\sigma > 0$, where $a > 1$. Precisely, we want to determine

$$R(\sigma) \ := \ \max_{x\in(0,1)} \frac{\beta_{a,a}}{\phi_{1/2,\sigma}}(x),$$

because for arbitrary Borel sets $S \subset \mathbb{R}$,

$$\mathrm{Beta}(a,a)(S) \ \le \ R(\sigma)\,\mathcal{N}(1/2,\sigma^2)(S)$$

and

$$\bigl|\mathrm{Beta}(a,a)(S) - \mathcal{N}(1/2,\sigma^2)(S)\bigr| \ \le \ 1 - R(\sigma)^{-1}.$$

Moreover, we want to find $\sigma > 0$ such that this quantity is minimal.

To determine $R(\sigma)$, note first that for fixed a and σ,

$$\begin{aligned}
\log \frac{\beta_{a,a}}{\phi_{1/2,\sigma}}(x) \ &= \ \log\sqrt{2\pi\sigma^2} - \log B(a,a) + \frac{(x-1/2)^2}{2\sigma^2} + (a-1)\log(x(1-x)) \\
&= \ \log\sqrt{2\pi\sigma^2} - \log B(a,a) + \frac{(x-1/2)^2}{2\sigma^2} + (a-1)\log\bigl(1/4-(x-1/2)^2\bigr) \\
&= \ \mathrm{const}(a,\sigma) + \frac{y}{8\sigma^2} + (a-1)\log(1-y),
\end{aligned}$$

where $y := (2x-1)^2 \in [0,1)$. Since

$$\frac{d}{dy}\Bigl(\frac{y}{8\sigma^2} + (a-1)\log(1-y)\Bigr) \ = \ \frac{1}{8\sigma^2} - \frac{a-1}{1-y},$$

the maximum of $\log(\beta_{a,a}/\phi_{1/2,\sigma})$ is attained at $x \in (0,1)$ such that $y = \bigl(1-8\sigma^2(a-1)\bigr)^+$, and the resulting value of $\log R(\sigma)$ is

$$\begin{aligned}
\log R(\sigma) \ &= \ \log\sqrt{2\pi} - \log B(a,a) + (a-1)\log(1/4) \\
&\quad + \ \log(\sigma^2)/2 + \bigl((8\sigma^2)^{-1} - a + 1\bigr)^+ + (a-1)\log\min\{8\sigma^2(a-1),1\} \\
&= \ \log\sqrt{2\pi} - \log B(a,a) - (2a-1/2)\log(2) \\
&\quad + \ \log(8\sigma^2)/2 + \bigl((8\sigma^2)^{-1} - a + 1\bigr)^+ + (a-1)\log\min\{8\sigma^2(a-1),1\}.
\end{aligned}$$

This is strictly monotone increasing in $8\sigma^2 \ge (a-1)^{-1}$, so we restrict our attention to values σ in $\bigl(0,(8(a-1))^{-1/2}\bigr]$. Then,

$$\begin{aligned}
\text{(5)} \qquad \log R(\sigma) \ &= \ \log\sqrt{2\pi} - \log B(a,a) - (2a-1/2)\log(2) \\
&\quad + \ (8\sigma^2)^{-1} + (a-1/2)\log(8\sigma^2) - a + 1 + (a-1)\log(a-1).
\end{aligned}$$

Note also the Stirling type approximation

$$\log\Gamma(y) \ = \ \log\sqrt{2\pi} + (y-1/2)\log(y) - y + r(y),$$

where $r(y) = 1/(12y) + O(y^{-2})$ as $y \to \infty$ (cf. Dümbgen et al. 2021, Lemma 10). Consequently,

$$\begin{aligned}\log\sqrt{2\pi} - \log B(a,a) &= \log\sqrt{2\pi} + \log\Gamma(2a) - 2\log\Gamma(a) \\ &= (2a-1/2)\log(2a) - 2(a-1/2)\log(a) + r(2a) - 2r(a) \\ &= (2a-1/2)\log(2) + \log(a)/2 + \tilde{r}(a),\end{aligned}$$

with $\tilde{r}(a) := \tilde{r}(2a) - 2\tilde{r}(a)$. This leads to

$$\begin{aligned}\log R(\sigma) &= \tilde{r}(a) + \log(a)/2 \\ &\quad + (8\sigma^2)^{-1} + (a-1/2)\log(8\sigma^2) - a + 1 + (a-1)\log(a-1).\end{aligned} \tag{6}$$

For the particular choice of σ, there are at least three possibilities:

Moment matching. A first candidate for σ would be the standard deviation of $\mathrm{Beta}(a,a)$,

$$\sigma_1(a) := (8(a+1/2))^{-1/2}.$$

Local density matching. Since $\log\beta_{a,a}(x) - \log\beta_{a,a}(1/2)$ equals $-4(a-1)(x-1/2)^2$ plus a remainder of order $O((x-1/2)^4)$ as $x \to 1/2$, another natural choice would be

$$\sigma_2(a) := (8(a-1))^{-1/2}.$$

Minimizing $R(\sigma)$. Note that $\log R(\sigma) = \mathrm{const}(a) + (a-1/2)\log(8\sigma^2) + (8\sigma^2)^{-1}$. Since

$$\frac{d}{dy}\big((a-1/2)\log(y) + y^{-1}\big) = \frac{a-1/2}{y} - \frac{1}{y^2} = \frac{(a-1/2)(y-(a-1/2)^{-1})}{y^2},$$

the optimal value of σ equals

$$\sigma_3(a) := (8(a-1/2))^{-1/2}.$$

Numerical example. Figure 3 shows for $a = 5$ the beta density $\beta_{a,a}$ and the Gaussian approximations $\phi_{1/2,\sigma}$, where $\sigma = \sigma_1(a), \sigma_2(a), \sigma_3(a)$. Figure 4 depicts the corresponding log-density ratios $\log(\beta_{a,a}/\phi_{1/2,\sigma})$. The values of $R(\sigma)$, rounded to four digits, are $R(\sigma_1(a)) = 1.1660$, $R(\sigma_2(a)) = 1.0905$ and $R(\sigma_3(a)) = 1.0582$.

Our specific values $\sigma_j(a)$ are of the type $\sigma = (8(a+\delta))^{-1/2}$ for some $\delta \ge -1$. The next lemma provides two important properties of the resulting value $\log R(\sigma)$.

Lemma 11. *Let $\sigma(a) := (8(a+\delta))^{-1/2}$ for $a > 1$ with a fixed number $\delta \ge -1$. Then $\log R(\sigma(a))$ is strictly decreasing in $a > 1$, and*

$$\log R(\sigma) = \frac{\delta(\delta+1) + 3/4}{2a} + O(a^{-2}).$$

For our specific standard deviations $\sigma_j(a)$ we obtain the limits

$$\lim_{a\to\infty} a\log R(\sigma_j(a)) = \begin{cases} 3/4 & \text{if } j = 1, \\ 3/8 & \text{if } j = 2, \\ 1/4 & \text{if } j = 3. \end{cases}$$

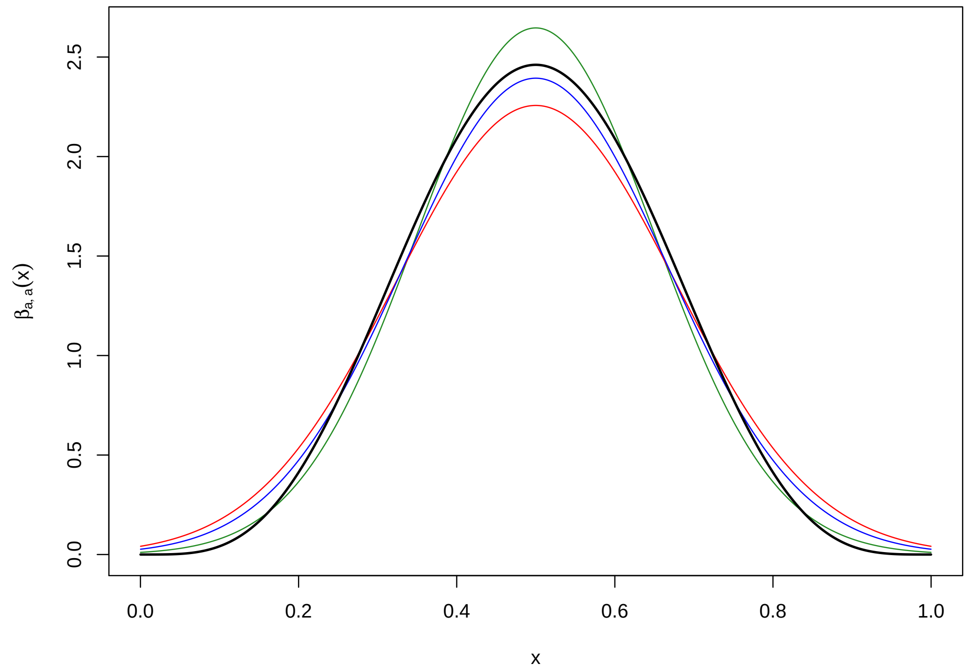

Figure 3: The density $\beta_{a,a}$ (black) for $a = 5$ and its Gaussian approximation $\phi_{1/2,\sigma}$ for $\sigma = (8(a+1/2))^{-1}$ (green), $\sigma = (8(a-1))^{-1/2}$ (red) and $\sigma = (8(a-1/2))^{-1/2}$ (blue).

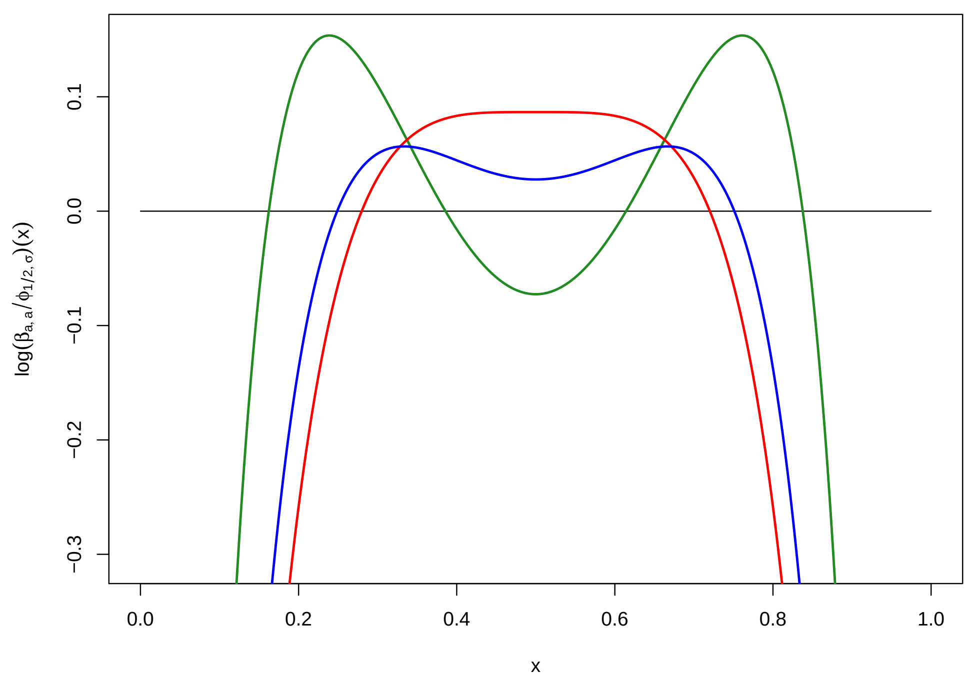

Figure 4: The log-density ratios $\log(\beta_{a,a}/\phi_{1/2,\sigma})$ for $a = 5$, where σ equals $\sigma_1(a)$ (green), $\sigma_2(a)$ (red) or $\sigma_3(a)$ (blue).

Remark 12. Similarly as in Section 4, we may conclude that for arbitrary $x > 1/2$ and $\sigma = (8(a+\delta))^{-1/2}$,

$$\bar{B}_{a,a}(x) \ \le\ R(\sigma)\Phi\Bigl(-\frac{x-1/2}{\sigma}\Bigr) \ \le\ \frac{R(\sigma)}{2}\exp\bigl(-4(a+\delta)(x-1/2)^2\bigr).$$

Even the latter bound is stronger than the bound $\exp\bigl(-4(a+1/2)(x-1/2)^2\bigr)$ by Marchal and Arbel (2017), as soon as $\delta \ge 0.5$ and $R(\sigma) \le 2$. For $\delta = 0.5$, this is the case for $a \ge 1.4$, and for $\delta = 1$, we only need $a \ge 1.9$.

6 Proofs

Proof of Theorem 1. Note that

$$Q_{a,b}(x) \ =\ \frac{a}{x^a}\int_0^x u^{a-1}(1-u)^{b-1}\,\mathrm{d}u \ =\ a\int_0^1 w^{a-1}(1-xw)^{b-1}\,\mathrm{d}w.$$

Since $\mathrm{d}^2(1-xw)^{b-1}/\mathrm{d}w^2 = (b-1)(b-2)x^2(1-xw)^{b-3}$, the function $w \mapsto (1-xw)^{b-1}$ is convex if $b \not\in (1,2)$ and concave if $b \in [1,2]$. Since $a\int_0^1 w^{a-1}\,\mathrm{d}w = 1$, it follows from Jensen's inequality that

$$\Bigl(a\int_0^1 w^{a-1}(1-xw)\,\mathrm{d}w\Bigr)^{b-1} \ =\ q_{a,b}^{(1)}(x)$$

is a lower bound for $Q_{a,b}$ if $b \not\in (1,2)$ and an upper bound if $b \in [1,2]$. To compare $Q_{a,b}$ with $q_{a,b}^{(2)}$, we use a well-known formula for linear interpolation of the function $w \mapsto (1-xw)^{b-1}$ with second derivative $(b-1)(b-2)x^2(1-xw)^{b-3}$ on $[0,1]$, namely,

$$(1-xw)^{b-1} \ =\ 1-w+w(1-x)^{b-1} - w(1-w)(b-1)(b-2)x^2(1-x\tilde{w})^{b-3}/2$$

for some $\tilde{w} = \tilde{w}(x,w) \in (0,1)$. Note that

$$(b-1)(b-2)(1-x\tilde{w})^{b-3} \ \begin{cases} \ge\ (b-1)(b-2) & \text{if } b \le 1, \\ \le\ (b-1)(b-2) & \text{if } b \in [1,2], \\ \ge\ (b-1)(b-2)(1-x)^{(b-3)^+} & \text{if } b \ge 2. \end{cases}$$

Hence, with $h_b(w) := 1-w+w(1-x)^{b-1} - w(1-w)(b-1)(b-2)x^2(1-x)^{(b-3)^+}/2$ we may conclude that

$$a\int_0^1 w^{a-1}h_b(w)\,\mathrm{d}w \ =\ q_{a,b}^{(2)}(x)$$

is an upper bound for $Q_{a,b}(x)$ if $b \not\in (1,2)$ and a lower bound if $b \in [1,2]$.

Concerning alternative bounds for $Q_{a,b}$, let $Q : [0,x_o] \to (0,\infty]$ be a continuous function for some $x_o \in (0,1]$. Viewing Q as a bound of $Q_{a,b}$, $H(x) = x^a Q(x)/[aB(a,b)]$ is a bound for $B_{a,b}(x)$. If Q is differentiable on $(0,x_o)$, then elementary calculus reveals that

$$H'(x) \ =\ \beta_{a,b}(x)J(x) \quad \text{with} \quad J(x) \ :=\ \frac{Q(x)+Q'(x)x/a}{(1-x)^{b-1}}.$$

If we can show that $J \geq 1$ (or $J \leq 1$) on $(0, x_o)$, we may conclude that $Q_{a,b} \leq Q$ (or $Q_{a,b} \geq Q$) on $[0, x_o]$. For instance, let $Q(x) := (1-x)^b(1+cx)$ for some $c > 0$ and $x \in [0,1]$. Then one can show that $J \leq 1$ on $[0,1]$, provided that $c \geq c_{a,b} = (a+b)/(a+1)$. This yields the lower bound for $Q_{a,b}$ in (1). Now, let $Q(x) := (1-x)^b/(1-cx)$ for some $c > 0$ and $0 \leq x \leq x_o := \min\{c^{-1}, 1\}$. For $0 < x < x_o$,

$$J(x) \;=\; 1 + \frac{x}{a(1-cx)}\Big(c(a+1) - (a+b) + (c-1)\frac{cx}{1-cx}\Big).$$

If $c < 1$, then the infimum of $c(a+1) - (a+b) + (c-1)cx/(1-cx)$ over all $x \in (0, x_o)$ equals $ca - (a+b) < 0$. If $c \geq 1$, that infimum equals $c(a+1) - (a+b) \geq 0$, provided that $c \geq (a+b)/(a+1)$. Consequently, $J \geq 1$ on $(0, x_o)$ if $c \geq \max\{c_\ell, 1\}$, and this yields the upper bound in (1) as well as the upper bound $q_{a,b}^{(3)}(x)$ for $Q_{a,b}(x)$ in case of $b \geq 1$.

It remains to verify the additional inequalities for $Q_{a,b}^L, Q_{a,b}^U$. Concerning the lower bound for $Q_{a,b}^L$, the inequality $q_{a,b}^{(1)}(x) > (1-x)^b(1 + c_{a,b}x)$ is equivalent to

$$\big(1 - x/(a+1-ax)\big)^{-b} \;>\; 1 + bx/(a+1) - a(a+b)x^2/(a+1)^2.$$

Indeed, by convexity of $(1-\cdot)^{-b}$, the left-hand side is larger than $1 + bx/(a+1-ax) > 1 + bx/(a+1)$. Now let $b \geq 1$. For $b \in [1,2]$, $q_{a,b}^{(2)}(x) \geq (a(1-x)^{b-1}+1)/(a+1)$, and the latter term is strictly larger than $(1-x)^b(1+c_{a,b}x)$ if and only if

$$(1-x)^{-(b-1)} \;>\; 1 + (b-1)x - (a+b)x^2.$$

Indeed, since $b - 1 \geq 0$, the left-hand side is not smaller than $1 + (b-1)x$.

Concerning the upper bound for $Q_{a,b}^U$, if $b \leq 1$, then $c_{a,b} \leq 1$ and $(1-x)^{b-1} \geq 1$, so $q_{a,b}^{(2)}(x) \leq (a(1-x)^{b-1}+1)/(a+1) \leq (1-x)^{b-1} = (1-x)^b/(1 - \max\{c_{a,b}, 1\}x)^+$. If $1 < b \leq 2$, then $c_{a,b} > 1$, and the inequality $q_{a,b}^{(1)}(x) < (1-x)^b/(1-c_{a,b}x)^+$ is equivalent to

$$(1 - (b-1)y)^+(1+y)^{b-1} \;<\; 1$$

with $y := (a+1)^{-1}x/(1-x) \in (0, \infty)$. By concavity of $(1+\cdot)^{b-1}$, $(1+y)^{b-1} \leq 1 + (b-1)y$, whence $(1-(b-1)y)^+(1+y)^{b-1} \leq (1-(b-1)^2y^2)^+ < 1$. $\square$

Proof of Corollary 4. Recall Stirling's approximation $\Gamma(c) = \sqrt{2\pi}c^{c-1/2}e^{-c}(1+o(1))$ as $c \to \infty$. This implies the following asymptotic expansions as $b \to \infty$:

$$\frac{1}{B(a,b)} \;=\; \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \;=\; \frac{b^a(1+a/b)^{a+b-1/2}e^{-a}(1+o(1))}{\Gamma(a)} \;=\; \frac{b^a(1+o(1))}{\Gamma(a)}.$$

Consequently,

$$G_a(x) \;=\; \lim_{b\to\infty} B_{a,b}(x/b) \;=\; \lim_{b\to\infty} \frac{(x/b)^a}{aB(a,b)}Q_{a,b}(x/b) \;=\; \frac{x^a}{a\Gamma(a)} \lim_{b\to\infty} Q_{a,b}(x/b).$$

Bounding $Q_{a,b}(x/b)$ in terms of $q_{a,b}^{(\ell)}(x/b)$, $1 \le \ell \le 3$, as in Theorem 1, the asserted bounds for $G_a(x)$ follow immediately from the following limits:

$$\begin{aligned}
q_{a,b}^{(1)}(x/b) &= \Big(1 - \frac{ax}{(a+1)b}\Big)^{b-1} \to e^{-ax/(a+1)}, \\
q_{a,b}^{(2)}(x/b) &= \frac{a(1-x/b)^{b-1}+1}{a+1} - \frac{a(b-1)(b-2)x^2(1-x/b)^{(b-3)^+}}{2b^2(a+1)(a+2)} \\
&\to \frac{ae^{-x}+1}{a+1} - \frac{ax^2e^{-x}}{2(a+1)(a+2)}, \\
q_{a,b}^{(3)}(x/b) &= \frac{(1-x/b)^b}{\big(1-[(a+b)/b]x/(a+1)\big)^+} \to \frac{e^{-x}}{(1-x/(a+1))^+}.
\end{aligned}$$

As to $\bar{G}_a(x)$, we write $\bar{G}_a(x) = \lim_{b\to\infty} \bar{B}_{a,b}(x/b) = \lim_{b\to\infty} B_{b,a}(1-x/b)$ and

$$B_{b,a}(1-x/b) = \frac{(1-x/b)^b}{bB(a,b)} Q_{b,a}(1-x/b) = \frac{e^{-x}b^{a-1}(1+o(1))}{\Gamma(a)} Q_{b,a}(1-x/b),$$

so $\bar{G}_a(x)$ is $e^{-x}/\Gamma(a)$ times $\lim_{b\to\infty} b^{a-1}Q_{b,a}(1-x/b)$. Bounding $Q_{b,1}(1-x/b)$ in terms of $q_{b,1}^{(\ell)}(1-x/b)$, $1 \le \ell \le 3$, as in Theorem 1, the asserted bounds for $\bar{G}_a(x)$ follow immediately from the following limits:

$$\begin{aligned}
b^{a-1}q_{b,a}^{(1)}(1-x/b) &= b^{a-1}\Big(\frac{x+1}{b+1}\Big)^{a-1} \to (x+1)^{a-1}, \\
b^{a-1}q_{b,a}^{(2)}(1-x/b) &= \frac{bx^{a-1}+b^{a-1}}{b+1} - \frac{b^a(a-1)(a-2)(1-x/b)^2(x/b)^{(a-3)^+}}{2(b+1)(b+2)} \\
&\to \begin{cases} x^{a-1} & \text{if } a < 2, \\ x+1 & \text{if } a = 2, \\ \infty & \text{if } a > 2, a \ne 3, \\ x^2+2x+2 & \text{if } a = 3, \end{cases} \\
b^{a-1}q_{b,a}^{(3)}(1-x/b) &= x^a \Big/ \Big(\frac{(a+b)x - b(a-1)}{b+1}\Big)^+ \to \frac{x^a}{(x-a+1)^+}.
\end{aligned}$$

$\square$

Proof of Theorem 5. Since $B_{a,b}(\cdot) = \bar{B}_{b,a}(1-\cdot)$ and $K(q,x) = K(1-q,1-x)$ for $q \in (0,1)$ and $x \in [0,1]$, it suffices to prove the result for $\bar{B}_{a,b}(x)$, $x \in [p_r, 1]$.

In case of $a = 1$, the asserted bounds are sharp, because $B(a,b) = 1/b$, $p_r = 0$ and $\bar{B}_{a,b}(x) = (1-x)^b$. In case of $a > 1$, the ratio

$$Q(x) := \frac{\bar{B}_{a,b}(x)}{x^{a-1}(1-x)^b} = \frac{B(a,b)^{-1}}{1-x}\int_x^1 \Big(\frac{u}{x}\Big)^{a-1}\Big(\frac{1-u}{1-x}\Big)^{b-1}\,\mathrm{d}u$$

is strictly decreasing in $x \in (0,1)$. Indeed, with $w(u) := (1-u)/(1-x) \in (0,1)$ for $u \in (x,1)$, we have $dw(u)/du = -1/(1-x)$, and $u = 1-(1-x)w(u)$, so

$$Q(x) = B(a,b)^{-1}\int_0^1 \Big(w + \frac{1-w}{x}\Big)^{a-1} w^{b-1}\,\mathrm{d}w,$$

which is strictly decreasing in $x \in (0,1)$ with limit $Q(1) = 1/[bB(a,b)]$. Consequently, for $0 < x_o \le x \le 1$,

$$Q(1) \ \le \ Q(x) \ \le \ Q(x_o).$$

Multiplying these inequalities with $x^{a-1}(1-x)^b$ and setting $x_o = p_r$ yields the asserted bounds for $\bar{B}_{a,b}(x)$. □

Proof of Corollary 9. For symmetry reasons, it suffices to derive the upper bound for $\bar{B}_{a,b}(x)$. It suffices to show that for $x \in [p,1]$,

$$(a+b-1)K(p_r,x) \ \ge \ \frac{(a+b+1)(x-p)^2}{2p(1-p) + (4/3)(1-2p)^+(x-p)}.$$

To simplify notation, we write $m := a+b$, $y := x-p \in [0,1-p]$ and $\delta := p-p_r = (1-p)/(m-1)$. Then it follows from the first inequality in (4) that

$$(a+b-1)K(p_r,x) \ \ge \ \frac{(m-1)(y+\delta)^2}{2(p+(2/3)y-\delta/3)(1-p-(2/3)y+\delta/3)},$$

and we want to show that this is greater than or equal to

$$\frac{(m+1)y^2}{2[p(1-p)+(2/3)(1-2p)y]}.$$

Note first that since $(m-1)\delta = (1-p)$,

$$\begin{aligned} (m-1)(y+\delta)^2 \ &= \ (m-1)y^2 + 2(1-p)y + (1-p)\delta \\ &= \ (m+1)y^2 + 2(1-p-y)y + (1-p)\delta \\ &> \ (m+1)y^2. \end{aligned} \tag{7}$$

Furthermore,

$$\begin{aligned} &(p+(2/3)y-\delta/3)(1-p-(2/3)y+\delta/3) \\ &\quad = \ p(1-p) + (2/3)(1-2p)y + (2p-1)\delta/3 - (2y-\delta)^2/9, \end{aligned} \tag{8}$$

and in case of $0 < p \le 1/2$, the right hand side is not larger than $p(1-p)+(2/3)(1-2p)y$. This proves our assertion in case of $0 < p \le 1/2$ already, and it remains to treat the case $1/2 < p < 1$. To this end, we have to show that the ratio of (7) and $(m+1)y^2$ is not smaller than the ratio of (8) and $p(1-p)+(2/3)(1-2p)y$. This assertion is equivalent to the inequality

$$\frac{2(1-p-y)y + (1-p)\delta}{(m+1)y^2} \ \ge \ \frac{(2p-1)\delta/3 - (2y-\delta)^2/9}{p(1-p)+(2/3)(1-2p)y}.$$

With $\lambda := (m-1)^{-1}$ and $z := y/(1-p) \in [0,1]$, the latter inequality is equivalent to

$$\frac{2(1-z)z+\lambda}{(\lambda^{-1}+2)z^2} \ \ge \ \frac{(2p-1)\lambda - (1-p)(2z-\lambda)^2/3}{3p+2(1-2p)z}.$$

Since the left-hand side is strictly positive and $3p + 2(1-2p)z \geq 3p + 2(1-2p) = 2-p$, it suffices to show that

$$\text{(9)} \qquad \frac{(1-p)(2z-\lambda)^2}{3(2-p)} + \frac{2(1/z-1)+\lambda/z^2}{1+2\lambda}\lambda \ \geq \ \frac{(2p-1)}{2-p}\lambda.$$

If $z \leq 2/3$, the second summand on the left-hand side is at least

$$\frac{2(3/2-1)+\lambda(3/2)^2}{1+2\lambda}\lambda \ = \ \frac{1+(9/4)\lambda}{1+2\lambda}\lambda \ > \ \lambda \ > \ \frac{2p-1}{2-p}\lambda.$$

Thus it suffices to consider $z \geq 2/3$. It follows from $1 \leq b = (1-p)m$ that $m \geq 1/(1-p)$, whence $\lambda \leq (1-p)/p$. Thus $2z - \lambda \geq 2z - (1-p)/p > 0$ and $(1-p) \geq p\lambda$. Consequently, it suffices to verify that

$$\text{(10)} \qquad \frac{p(2z-(1-p)/p)^2}{3(2-p)} + \frac{2(1/z-1)+\lambda/z^2}{(1+2\lambda)} \ \geq \ \frac{(2p-1)}{2-p}.$$

The second summand on the left-hand side equals

$$\frac{1}{2z^2}\frac{4z(1-z)+2\lambda}{1+2\lambda} \ = \ \frac{1}{2z^2}\Big(1 - \frac{1-4z(1-z)}{1+2\lambda}\Big),$$

an increasing function of $\lambda > 0$ for any fixed $z \in (0,1]$. Consequently, inequality (10) would be a consequence of

$$\frac{p(2z-(1-p)/p)^2}{3(2-p)} + 2(1/z-1) \ \geq \ \frac{(2p-1)}{2-p}.$$

Since $1/z - 1 = (1-z)/z \geq 1-z$, it even suffices to show that

$$\text{(11)} \qquad (p/3)(2z-(1-p)/p)^2 + 2(2-p)(1-z) \ \geq \ 2p-1.$$

The minimiser of the left-hand side, as a function of $z \in \mathbb{R}$, is given by

$$z_o \ = \ 2/p - 5/4 \ > \ 3/4.$$

If $p \leq 8/9$, then $z_o \geq 1$, so it suffices to verify (11) for $z = 1$. Indeed,

$$(p/3)(2-(1-p)/p)^2) \ = \ (4/3)(2p-1) + (1-p)^2/(3p) \ > \ 2p-1.$$

If $8/9 \leq p < 1$, then (11) is equivalent to

$$2p-1 \ \leq \ (p/3)(2z_o-(1-p)/p)^2 + 2(2-p)(1-z_o) \ = \ 10 - 5/p - (15/4)p.$$

But this inequality is equivalent to

$$\text{(12)} \qquad (23/4)p + 5/p \ \leq \ 11.$$

The left-hand side is convex in p, so it suffices to verify it for $p = 8/9$ and $p = 1$. The left-hand side of (12) equals $46/9 + 45/8 = 10 + 1/9 + 5/8 < 11$ if $p = 8/9$, and $10 + 3/4 < 11$ if $p = 1$. This concludes our proof of Corollary 9. $\square$

Proof of Lemma 11. At first we analyze $\tilde{r}(a)$. We use Binet's fomula $r(y) = \int_0^\infty e^{-yt} w(t)\,\mathrm{d}t$ with a certain function w satisfying $12^{-1} e^{-t/12} < w(t) < 12^{-1}$, see Dümbgen et al. (2021, Lemma 10). Consequently, $2r(a) - r(2a) = \int_0^\infty (2e^{-at} - e^{-2at}) w(t)\,\mathrm{d}t$, and since $2e^{-at} - e^{-2at} = e^{-at}(2 - e^{-at}) > 0$, we conclude that

$$2r(a) - r(2a) \begin{cases} < \dfrac{1}{12} \displaystyle\int_0^\infty (2e^{-at} - e^{-2at})\,\mathrm{d}t \ = \ \dfrac{1}{8a}, \\ > \dfrac{1}{12} \displaystyle\int_0^\infty (2e^{-(a+1/12)t} - e^{-(2a+1/12)t})\,\mathrm{d}t \ = \ \dfrac{a + 1/36}{8(a+1/12)(a+1/24)}. \end{cases}$$

In particular, as $a \to \infty$,

$$\tilde{r}(a) \ = \ -\frac{1}{8a} + O(a^{-2}). \tag{13}$$

Moreover,

$$\frac{d}{da}\tilde{r}(a) \ = \ 2\int_0^\infty t(e^{-at} - e^{-2at}) w(t)\,\mathrm{d}t \ < \ \frac{1}{6}\int_0^\infty t(e^{-at} - e^{-2at})\,\mathrm{d}t \ = \ \frac{1}{8a^2}. \tag{14}$$

Next we verify that $\log R(\sigma(a))$ is strictly decreasing in $a > 1$. It follows from representation (6) that

$$\begin{aligned} \log R(\sigma(a)) \ &= \tilde{r}(a) + \log(a)/2 \\ &\qquad + \ \delta - (a - 1/2)\log(a + \delta) + 1 + (a-1)\log(a-1). \end{aligned} \tag{15}$$

According to (14), the derivative of this with respect to a is not greater than

$$\frac{1}{8a^2} + \frac{1}{2a} - \frac{a - 1/2}{a + \delta} - \log(a+\delta) + \log(a-1) + 1.$$

For fixed $a > 1$, the derivative of this bound with respect to $\delta \ge 1$ equals $-(\delta + 1/2)/(a+\delta)^2$, so it is maximal for $\delta = -1/2$. This leads to

$$\begin{aligned} \frac{d}{da}\log R(\sigma(a)) \ &\le \ \frac{1}{8a^2} + \frac{1}{2a} + \log\Bigl(\frac{a-1}{a-1/2}\Bigr) \ = \ \frac{1}{8a^2} + \frac{1}{2a} + \log\Bigl(1 - \frac{1}{2(a-1/2)}\Bigr) \\ &< \ \frac{1}{8a^2} + \frac{1}{2a} + \log\Bigl(1 - \frac{1}{2a}\Bigr) \ = \ -\sum_{k\ge 3}(2a)^{-k}/k \ < \ 0. \end{aligned}$$

It remains to prove the expansion of $\log R(\sigma(a)))$ as $a \to \infty$. To this end, we rewrite (15) as

$$\log R(\sigma(a)) \ = \ \tilde{r}(a) + \delta - (a - 1/2)\log(1 + \delta/a) + (a-1)\log(1 - 1/a).$$

Since $\log(1+y) = y + O(y^2) = y - y^2/2 + O(y^3)$ as $y \to 0$,

$$\begin{aligned} \delta - (a-1/2)\log(1+\delta/a) \ &= \ \delta - \frac{(a-1/2)\delta}{a} + \frac{(a-1/2)\delta^2}{2a^2} + O(a^{-2}) \\ &= \ \frac{\delta(\delta+1)}{2a} + O(a^{-2}), \\ 1 + (a-1)\log(1-1/a) \ &= \ 1 - \frac{a-1}{a} - \frac{a-1}{2a^2} + O(a^{-2}) \\ &= \ \frac{1}{2a} + O(a^{-2}). \end{aligned}$$

Combining these expansions with (13) leads to the desired expansion of $\log R(\sigma(a))$. □

Acknowledgements. We are grateful to Maciej Skorski and two anonymous referees for constructive comments. This work was supported by the Swiss National Science Foundation.

References

[1] DIMITRIADIS, T., DÜMBGEN, L., HENZI, A., PUKE, M. and ZIEGEL, J. (2022). Honest calibration assessment for binary outcome predictions. *Biometrika*, to appear. (arXiv:2203.04065)

[2] DÜMBGEN, L. (1998). New goodness-of-fit tests and their application to nonparametric confidence sets. *Ann. Statist.* **26** 288–314.

[3] DÜMBGEN, L., SAMWORTH, R. J. and WELLNER, J. A. (2021). Bounding distributional errors via density ratios. *Bernoulli* **27** 818 – 852.

[4] DÜMBGEN, L. and WELLNER, J. A. (2022). A New Approach to Tests and Confidence Bands for Distribution Functions. *Ann. Statist.*, to appear. (arXiv:1402.2918).

[5] GROENEVELD, R. A. and MEEDEN, G. (1977). The mode, median, and mean inequality. *Amer. Statist.* **31** 120–121.

[6] GUPTA, A. K. and NADARAJAH, S. (eds.) (2004). *Handbook of beta distribution and its applications*, vol. 174 of *Statistics: Textbooks and Monographs*. Marcel Dekker, Inc., New York.

[7] MARCHAL, O. and ARBEL, J. (2017). On the sub-Gaussianity of the beta and Dirichlet distributions. *Electron. Commun. Probab.* **22** Paper No. 54, 14.

[8] NEUMAN, E. (2013). Inequalities and bounds for the incomplete gamma function. *Res. Math.* **63** 1209–1214.

[9] PINELIS, I. (2020). Exact lower and upper bounds for the incomplete gamma function. *arXiv e-prints* arXiv:2005.06384.

[10] QI, F. and MEI, J.-Q. (1999). Some inequalities of the incomplete gamma and related functions. *Z. Anal. Anwendungen* **18**(3) 793–799.

[11] SEGURA, J. (2014). Monotonicity properties and bounds for the chi-square and gamma distributions. *Appl. Math. Comput.* **246** 399–415.

[12] SEGURA, J. (2016). Sharp bounds for cumulative distribution functions. *J. Math. Anal. Appl.* **436** 748–763.

[13] SKORSKI, M. (2021). Bernstein-type bounds for beta distributions. *arXiv e-prints* arXiv:2101.02094v3.