Appendix to Quicksort Asymptotics

James Allen Fill[1]

Department of Mathematical Sciences

The Johns Hopkins University

jimfill@jhu.edu and http://www.mts.jhu.edu/~fill/

and

Svante Janson

Department of Mathematics

Uppsala University

svante.janson@math.uu.se and http://www.math.uu.se/~svante/

ABSTRACT

This appendix to [2] contains a proof of the improved estimates in Remark 7.3 of that paper for the moment generating function of the (normalized) number of comparisons in Quicksort.

AMS 2000 *subject classifications.* Primary 68W40; secondary 68P10, 60F05, 60E10.

Date. May 22, 2001.

[1]Research supported by NSF grant DMS–9803780, and by The Johns Hopkins University's Acheson J. Duncan Fund for the Advancement of Research in Statistics.

This is an appendix to [2], to which we refer for background and notation. The theorem, lemmas, and equations in this appendix are labelled by A.1, etc.; labels with pure numbers refer to [2].

The purpose of this appendix is to provide a proof of the following estimates stated in Remark 7.3 of [2].

Theorem A.1. *Let $L_0 \doteq 5.018$ be the largest root of $e^L = 6L^2$. Then, for all $n \geq 0$,*

$$\mathbf{E}\, e^{\lambda Y_n} \leq \begin{cases} e^{1.34\lambda^2}, & \lambda \leq -0.58, \\ e^{0.5\lambda^2}, & -0.58 \leq \lambda \leq 0, \\ e^{\lambda^2}, & 0 \leq \lambda \leq 0.42, \\ e^{12\lambda^2}, & 0.42 \leq \lambda \leq L_0, \\ e^{2e^{\lambda}}, & L_0 \leq \lambda. \end{cases}$$

In particular, $\mathbf{E}\, e^{\lambda Y_n} \leq \exp\left(\max\left(12\lambda^2, 2e^{\lambda}\right)\right)$ *for all* $\lambda \in \mathbf{R}$.

The proof below follows closely the corresponding proof in [1], where we obtained by the method of Rösler [3] (with some refinements) explicit estimates for the moment generating function of the limit variable Y. In this appendix we treat instead the normalized number of comparisons Y_n for finite n. In the present case, some estimates involving $C_n(i)$, stated as lemmas below, become harder than the corresponding estimates in [1] where the limit as $n \to \infty$ is treated. Note that the bound in Theorem A.1 is the same as the one obtained for $\mathbf{E}\, e^{\lambda Y}$ in [1] for $\lambda \geq 0$, but slightly weaker for $\lambda < 0$ (or rather for $\lambda < -0.58$). (It seems likely that with further effort one could show that the bounds in [1] for $\mathbf{E}\, e^{\lambda Y}$ hold also for $\mathbf{E}\, e^{\lambda Y_n}$ for all λ and n, but this is still an open problem.)

In order to obtain good estimates we use extensive numerical calculations for small n to supplement our analytical estimates; we could do without these numerical calculations at the cost of increasing the constants in the exponents in the theorem. [All numerical calculations have been verified independently by the two authors, the (alphabetically) first using `Mathematica` and the second using `Maple`.]

We begin with some estimates of $C_n(i)$.

Lemma A.2. *The sequence $(\mu_n)_{n\geq 0}$ is nondecreasing and convex.*

Proof. By (1.7) and (1.8), for $n \geq 0$,

$$\begin{aligned} \mu_{n+1} - \mu_n &= 2(n+2)H_{n+1} - 4(n+1) - [2(n+1)H_{n+1} - 4(n+1) + 2] \\ &= 2H_{n+1} - 2, \end{aligned} \tag{A.1}$$

which is nonnegative and increasing. □

Lemma A.3. *For every $n \geq 1$, the sequence $(C_n(i))_{1\leq i\leq n}$ is convex. Its maximum is $C_n(1) = C_n(n)$ and its minimum is $C_n(\lfloor (n+1)/2 \rfloor) = C_n(\lceil (n+1)/2 \rceil)$.*

Proof. The definition (1.5) and Lemma A.2 show that $(C_n(i))_{1\leq i\leq n}$ is convex. Moreover, $C_n(i) \equiv C_n(n+1-i)$, and the result follows. □

Lemma A.4. *If* $1 \le i \le n$*, then*

$$-\eta \le C_n(i) \le 1,$$

where $\eta := 2\ln 2 - 1 \doteq 0.3863$*.*

Proof. By Lemma A.3 and (1.5)

$$C_n(i) \le C_n(1) = \tfrac{1}{n}(n-1+\mu_0+\mu_{n-1}-\mu_n) \le \tfrac{n-1}{n} \le 1,$$

because $\mu_0 = 0$ and $\mu_{n-1} \le \mu_n$ by Lemma A.2.

For the lower bound we first consider n odd, $n = 2m-1$ with $m \ge 1$. By Lemma A.3, (1.5), and (1.8),

$$\begin{aligned} C_{2m-1}(i) &\ge C_{2m-1}(m) = \tfrac{1}{2m-1}(2m-2+\mu_{m-1}+\mu_{m-1}-\mu_{2m-1}) \\ &= \tfrac{1}{2m-1}\left[2m-2+2(2mH_m-4m+2)-(4mH_{2m}-8m+2)\right] \\ &= \tfrac{2m}{2m-1}(1+2H_m-2H_{2m}). \end{aligned} \tag{A.2}$$

Note that for $k \ge 2$ we have $\ln k - \ln(k-1) = -\ln(1-\frac{1}{k}) > \frac{1}{k} + \frac{1}{2k^2}$, and thus

$$\begin{aligned} \delta_m &:= 2\ln 2 + 2(H_m - H_{2m}) = 2\sum_{k=m+1}^{2m}\left(\ln k - \ln(k-1) - \frac{1}{k}\right) \\ &> \sum_{k=m+1}^{2m}\frac{2}{2k^2} > \sum_{k=m+1}^{2m}\left(\frac{1}{k}-\frac{1}{k+1}\right) = \frac{1}{m+1} - \frac{1}{2m+1} = \frac{m}{(m+1)(2m+1)}. \end{aligned}$$

Hence, if $m \ge 2$, then

$$2m\delta_m = \frac{2m^2}{(m+1)(2m+1)} \ge \frac{8}{15} > \eta,$$

while if $m = 1$, then $\delta_1 = \eta$. Therefore,

$$-C_{2m-1}(i) \le -C_{2m-1}(m) = \frac{2m}{2m-1}(-1+2\ln 2-\delta_m) < \frac{2m}{2m-1}\left(\eta - \frac{\eta}{2m}\right) = \eta.$$

If $n = 2m$ is even, then Lemma A.3, (1.5), (1.7), and (1.8) similarly yield

$$\begin{aligned} C_{2m}(i) &\ge C_{2m}(m) = \tfrac{1}{2m}(2m-1+\mu_{m-1}+\mu_m-\mu_{2m}) \\ &= \tfrac{1}{2m}\left[2m-1+2mH_m-4m+2+2(m+1)H_m-4m\right. \\ &\qquad \left. -(2(2m+1)H_{2m}-8m)\right] \\ &= \tfrac{2m+1}{2m}(1+2H_m-2H_{2m}). \end{aligned}$$

Comparing with (A.2) we find by the estimate above $|C_{2m}(m)| < |C_{2m-1}(m)| < \eta$, and the result follows. ∎

Lemma A.5. *For $n \geq 1$ and $U_n \sim \mathrm{unif}\{1, \ldots, n\}$, the sequence $\mathbf{E}\, C_n(U_n)^2$ is strictly increasing, and therefore*

$$\mathbf{E}\, C_n(U_n)^2 = \frac{1}{n} \sum_{i=1}^{n} C_n(i)^2 < \mathbf{E}\, C(U)^2 = \sigma^2/3 \doteq 0.140.$$

Proof. We could use Lemma 2.2, Minkowski's inequality, and numerical calculations by computer to verify $\mathbf{E}\, C_n(U_n) < 0.15$, but we can do slightly better. Indeed, from the results in Section 1, one obtains the formula

$$\mathbf{E}\, C_n(U_n)^2 = \tfrac{7}{3}\left(1 + \tfrac{1}{n}\right)^2 - \tfrac{4}{3}\left(1 + \tfrac{2}{n}\right)\left(1 + \tfrac{1}{n}\right) H_n^{(2)} - \tfrac{4}{3} n^{-3} H_n, \qquad n \geq 1.$$

From this expression it is simple (if somewhat laborious) to prove increasingness. Finally, the limiting value of $\mathbf{E}\, C_n(U_n)^2$ is $\mathbf{E}\, C(U)^2 = \sigma^2/3$. □

Lemma A.6. *For $1 \leq i \leq n$,*

$$C_n(i) - 2\eta \left[\left(\frac{i-1}{n} \right)^2 + \left(\frac{n-i}{n} \right)^2 - 1 \right] \geq 0.$$

Proof. Fix n and denote the left-hand side by x_i. By (1.5) and (A.1), for $1 \leq i \leq n-1$ we have

$$\begin{aligned} n^2(x_{i+1} - x_i) &= n(\mu_i - \mu_{i-1} + \mu_{n-i-1} - \mu_{n-i}) \\ &\qquad + 2\eta\left[(i-1)^2 - i^2 + (n-i)^2 - (n-i-1)^2\right] \\ &= 2nH_i - 2nH_{n-i} + 2\eta(2n - 4i) \end{aligned}$$

and thus, for $1 \leq i \leq n-2$,

$$\begin{aligned} n^2(x_{i+2} - 2x_{i+1} + x_i) &= \frac{2n}{i+1} + \frac{2n}{n-i} - 8\eta = \frac{2n(n+1)}{(i+1)(n-i)} - 8\eta \geq \frac{2n(n+1)}{[(n+1)/2]^2} - 8\eta \\ &= \frac{8n}{n+1} - 8\eta \geq 4 - 8\eta > 0. \end{aligned}$$

Hence $(x_i)_{1 \leq i \leq n}$ is convex. Moreover $x_i = x_{n+1-i}$, and thus the minimum is x_{i_0} with $i_0 = \lfloor (n+1)/2 \rfloor$. Since $i_0 - 1 \leq n/2 \leq i_0$,

$$2\eta \left[\left(\frac{i_0 - 1}{n} \right)^2 + \left(\frac{n - i_0}{n} \right)^2 - 1 \right] \leq 2\eta \left(\frac{1}{4} + \frac{1}{4} - 1 \right) = -\eta \leq C_n(i_0)$$

by Lemma A.4. Hence $x_{i_0} \geq 0$ and the result follows. □

Lemma A.7. *If $1 \leq i \leq n$ and $(i-1)/n \leq u \leq i/n$, then*

$$u(1-u)C_n(i) \leq 0.05.$$

Proof. The left-hand side is not changed if we replace i by $n+1-i$ and u by $1-u$; hence we may assume that $i \le (n+1)/2$. Moreover if n is odd and $i=(n+1)/2$, then, by Lemma A.3, $C_n(i) = \min_j C_n(j)$, and since $\mathbf{E}\, C_n(U_n) = 0$ when $U_n \sim \text{unif}\{1,\dots,n\}$, $C_n(i) \le 0$ and the inequality is trivial.

We may thus assume $i \le n/2$. Since $u(1-u)$ is increasing on $[0,1/2]$, we may further assume $u = i/n$. Then, by (2.25),

$$u(1-u)C_n(i) \le u(1-u)\left(C(u) + \tfrac{3}{n}\right) \le u(1-u)C(u) + \tfrac{3}{4n}.$$

As stated in [1], it can easily be checked numerically that $\max_{0\le u\le 1} u(1-u)C(u) < 0.033$, and thus $u(1-u)C_n(i) < 0.05$ follows for $n \ge 45$. The cases $1 \le i \le n \le 44$ are verified numerically. (The maximum value is $591/12005 \doteq 0.0492$, obtained for $n=7$ and $i = 1$ or 7.) ☐

Proof of Theorem A.1. Let $U \sim \text{unif}(0,1)$ and, for $n \ge 1$, $K \ge 0$, $\lambda \in \mathbf{R}$,

$$\begin{aligned}
U_n &:= \lceil nU \rceil \sim \text{unif}\{1,\dots,n\},\\
W_n &:= \left(\frac{U_n-1}{n}\right)^2 + \left(\frac{n-U_n}{n}\right)^2 - 1 \le U^2 + (1-U)^2 - 1 = -2U(1-U),\\
f^*_{n,K}(\lambda) &:= \mathbf{E}\,\exp\left(\lambda C_n(U_n) + K\lambda^2 W_n\right),\\
f_{n,K}(\lambda) &:= \mathbf{E}\,\exp\left(\lambda C_n(U_n) - 2K\lambda^2 U(1-U)\right);
\end{aligned}$$

note that $f^*_{n,K}(\lambda) \le f_{n,K}(\lambda)$.

Suppose now that we have found positive numbers K and L such that

$$f^*_{n,K}(\lambda) \le 1, \qquad n \ge 1, \quad \lambda \in [0,L]. \tag{A.3}$$

Then, by induction, for every $n \ge 0$,

$$\mathbf{E}\, e^{\lambda Y_n} \le e^{K\lambda^2}, \qquad \lambda \in [0,L]. \tag{A.4}$$

Indeed, (A.4) is trivial for $n=0$, and if $n \ge 1$ and $\mathbf{E}\, e^{\lambda Y_m} \le e^{K\lambda^2}$ for $m \le n-1$ and $\lambda \in [0,L]$, then by the recursion (1.4), for $\lambda \in [0,L]$,

$$\begin{aligned}
\mathbf{E}\, e^{\lambda Y_n} &= \frac{1}{n}\sum_{i=1}^n \mathbf{E}\,\exp\left[\lambda\left\{\frac{i-1}{n}Y_{i-1} + \frac{n-i}{n}Y^*_{n-i} + C_n(i)\right\}\right]\\
&= \frac{1}{n}\sum_{i=1}^n \exp[\lambda C_n(i)]\left(\mathbf{E}\,\exp\left[\lambda\frac{i-1}{n}Y_{i-1}\right]\right)\left(\mathbf{E}\,\exp\left[\lambda\frac{n-i}{n}Y_{n-i}\right]\right)\\
&\le \frac{1}{n}\sum_{i=1}^n \exp[\lambda C_n(i)]\exp\left\{K\lambda^2\left[\left(\frac{i-1}{n}\right)^2 + \left(\frac{n-i}{n}\right)^2\right]\right\}\\
&= \mathbf{E}\,\exp\left[\lambda C_n(U_n) + K\lambda^2(W_n+1)\right]\\
&= e^{K\lambda^2} f^*_{n,K}(\lambda) \le e^{K\lambda^2}.
\end{aligned}$$

Similarly, if $f^*_{n,K}(\lambda) \le 1$ for every $n \ge 1$ and $\lambda \in [-L,0]$, then $\mathbf{E}\, e^{\lambda Y_n} \le e^{K\lambda^2}$ for every $n \ge 1$ and $\lambda \in [-L,0]$.

Thus our goal is to show $f^*_{n,K}(\lambda) \le 1$ for suitable K and λ; since $f^*_{n,K}(\lambda) \le f_{n,K}(\lambda)$, it suffices to show $f_{n,K}(\lambda) \le 1$. We follow the argument in [1], omitting many details which remain the same.

First, a Taylor expansion yields, using Lemma A.5, for $0 \le \lambda \le L$,

$$f_{n,K}(\lambda) \le 1 + \frac{1}{6}\lambda^2 \left(\sigma^2 - 2K + L \sup_{0 \le \lambda \le L} f'''_{n,K}(\lambda) \right). \tag{A.5}$$

Moreover,

$$\begin{aligned} f'''_{n,K}(\lambda) &= \mathbf{E}\Big[\Big((C_n(U_n) - 4K\lambda U(1-U))^3 - 12KU(1-U)\left(C_n(U_n) - 4K\lambda U(1-U)\right)\Big) \\ &\qquad \times \exp\left(\lambda C_n(U_n) - 2K\lambda^2 U(1-U)\right)\Big]. \end{aligned} \tag{A.6}$$

Using Lemma A.4, it follows as in [1] that

$$L \sup_{0 \le \lambda \le L} f'''_{n,K}(\lambda) \le L(3K\eta + 3K^2 L)e^L.$$

It is readily checked that for $K = 1$ and $L = 0.42$, this is less than $1.547 < 2K - \sigma^2$, so (A.5) shows that $f_{n,1}(\lambda) \le 1$ for $0 \le \lambda \le 0.42$. Hence (A.3) and thus (A.4) hold with $K = 1$ and $L = 0.42$.

For larger L we use again Lemma A.4 to obtain

$$f_{n,K}(\lambda) \le e^{|\lambda|} \mathbf{E}\, e^{-2K\lambda^2 U(1-U)}.$$

It is shown in [1] that the right-hand side is at most

$$g_K(\lambda) := e^{|\lambda|} \left[1 - \exp(-K\lambda^2/2)\right] / (K\lambda^2/2),$$

and further that $g_K(\lambda) < 1$ if $K = 12$ and $0.42 \le \lambda \le 2$, or if $K = 2L^{-2}e^L$ and $2 \le \lambda \le L$. It follows that (A.3) and (A.4) hold for any $L > 0$ and $K = \max\left(12, 2L^{-2}e^L\right)$.

For $-L \le \lambda \le 0$, a Taylor expansion yields [cf. (A.5)]

$$f_{n,K}(\lambda) \le 1 + \frac{1}{6}\lambda^2 \left(\sigma^2 - 2K + L \sup_{-L < \lambda \le 0} \left(-f'''_{n,K}(\lambda)\right) \right). \tag{A.7}$$

Moreover, from (A.6) and Lemmas A.4 and A.7, for $-L \le \lambda \le 0$ we have

$$f'''_{n,K}(\lambda) \ge (-\eta^3 - 12K \cdot 0.05 - 3K^2 L)e^{\eta L}.$$

Taking $K = 0.5$ and $L = 0.58$, we find

$$L \sup_{-L \le \lambda \le 0} \left(-f'''_{n,K}(\lambda)\right) < 0.576 < 2K - \sigma^2,$$

and thus by (A.7)

$$f_{n,0.5}(\lambda) \le 1, \qquad -0.58 \le \lambda \le 0. \tag{A.8}$$

Finally, for $\lambda \le -0.58$ we take $K = 2\eta/0.58 < 1.34$. Then $|K\lambda| \ge 2\eta$, and thus, using Lemma A.6,

$$\begin{aligned} \lambda C_n(U_n) + K\lambda^2 W_n &\le \lambda C_n(U_n) + 2\eta|\lambda| W_n \\ &= -|\lambda| \left[C_n(U_n) - 2\eta \left(\left(\frac{U_n - 1}{n} \right)^2 + \left(\frac{n - U_n}{n} \right)^2 - 1 \right) \right] \le 0. \end{aligned}$$

Hence $f^*_{n,K}(-\lambda) \le 1$. (This time we thus use $f^*_{n,K}$ instead of $f_{n,K}$.) Combined with (A.8), this shows that $f^*_{n,1.34}(\lambda) \le 1$ for all $\lambda \le 0$, which completes the proof. □

References

[1] Fill, J. A. and Janson, S. Approximating the limiting Quicksort distribution. Technical Report #615, Department of Mathematical Sciences, The Johns Hopkins University. Preprint available from `http://www.mts.jhu.edu/~fill/` or `http://www.math.uu.se/~svante/`.

[2] Fill, J. A. and Janson, S. Quicksort asymptotics. Technical Report #597, Department of Mathematical Sciences, The Johns Hopkins University. Preprint available from `http://www.mts.jhu.edu/~fill/` or `http://www.math.uu.se/~svante/`.

[3] Rösler, U. A limit theorem for 'Quicksort'. *RAIRO Inform. Théor. Appl.* **25** (1991), 85–100.