Bounding the Maximum of Dependent Random Variables

J.A. Hartigan

Yale University
e-mail: john.hartigan@yale.edu

Abstract: Abstract: Let M_n be the maximum of n zero-mean gaussian variables $X_1,..,X_n$ with covariance matrix of minimum eigenvalue λ and maximum eigenvalue Λ. Then, for $n \geq 70$,

$$\Pr\{M_n \geq \lambda\,(2\log n - 2.5 - \log(2\log n - 2.5))^{\frac{1}{2}} - .68\Lambda\} \geq \frac{1}{2}.$$

Bounds are also given for tail probabilities other than $\frac{1}{2}$. Upper bounds are given for tail probabilities of the maximum of dependent identically distributed variables. As an application, the maximum of purely non-deterministic stationary Gaussian processes is shown to have the same first order asymptotic behaviour as the maximum of independent gaussian processes.

MSC 2010 subject classifications: 60E15.
Keywords and phrases: maxima, tail probabilities.

1. Introduction.

The asymptotic behaviour of the maximum M_n of n i.i.d.(independent and identically distributed) random variables with continuous distribution function F is well known following Gumbel[2], namely that

$$-\log n - \log\left(1 - F(M_n)\right) \to G \text{ in distribution,}$$

where the Gumbel variable $G = -\log(-\log U)$ for U uniform.

When the variables are not independent, but identically distributed, upper bounds for the tail probabilities of M_n are available through

$$\Pr\{M_n \geq A\} \leq n\left(1 - F(A)\right),$$

but lower bounds are rare.

In [1], Berman examines conditions under which the maximum of a stationary process has a limiting distribution.

If the process is exchangeable, then the variables are i.i.d conditional on some tail process, and the limiting maximum is a distributed as mixture of limiting distributions of iid maxima conditioned on the the tail process. For example, if the process is gaussian, the limiting distribution will be that of $Y + M_n$ where Y is gaussian, and M_n is the maximum of n i.i.d guassians independent of Y.

For stationary Gaussian processes with correlations satisfying

$$\lim_{|i-j|\to\infty} \rho(X_i, X_j)\log|i-j| = 0,$$

the maximum behaves asymptotically like the maximum of i.i.d variables. Berman also presents conditions under which the maximum of a Markov chain has a limiting distribution.

Resnick[3] shows that if the variables form a markov chain, the limiting distribution of the maximum may be reduced to the distribution of the maximum of a set of i.i.d random variables.

The principal result here provides lower bounds for the tail probabilities of the maximum of dependent gaussian variables.

2. General bounds

We will use $\Pr X$ to denote the expectation of the random variable X, and $\{S\}$ to denote the function that is 1 when S is true, and 0 when S is false.

Theorem 2.1. *Let M_n denote the maximum of n random variables $X_1, ..X_n$ each with continuous distribution function F . Then, for each n, there exists an exponential variable E with*

$$-\log n - \log\left(1 - F(M_n)\right) \leq E. \tag{2.1}$$

Proof: Let F_n denote the distribution function of M_n. Then

$$\begin{aligned} \{M_n > A\} &\leq \sum_{i=1}^{n} \{X_i > A\}, \\ 1 - F_n(A) &\leq n\left(1 - F(A)\right), \\ -\log n - \log\left(1 - F(M_n)\right) &\leq E, \end{aligned}$$

since $1 - F_n(M_n) \sim \exp(-E)$.

Theorem 2.2. *Let M_n denote the maximum of n independent random variables each with continuous distribution function F . Then the function*

$$G(M_n) = -\log(-n \log F(M_n)),$$

has a Gumbel distribution and,

$$G \leq -\log\left(n(1 - F(M_n))\right) \leq G + \exp(-G)/n.$$

Proof: Since $F(M_n)$ is the maximum of n independent uniforms, $F(M_n) \sim U^{\frac{1}{n}}$, so $G = -\log(-n \log F(M_n))$ is Gumbel. Then

$$\begin{aligned} 1 - F(M_n) &= 1 - e^{-\frac{1}{n}\exp(-G)}, \\ \frac{1}{n}\exp(-G)/(1 + \frac{1}{n}\exp(-G)) &\leq 1 - e^{-\frac{1}{n}\exp(-G)} \leq \frac{1}{n}\exp(-G), \\ G \leq -\log\left(n(1 - F(M_n))\right) &\leq G + \log(1 + \frac{1}{n}\exp(-G)), \\ G \leq -\log\left(n(1 - F(M_n))\right) &\leq G + \frac{1}{n}\exp(-G). \end{aligned}$$

It follows that the limiting distribution of $-\log n - \log(1 - F(M_n))$ is the Gumbel distribution. Note that E and G are very close in their tail distributions, so there is not much difference between the upper bounds in the independent and dependent cases.

3. Gaussian bounds

In the Gaussian case, we invert the standard tail bounds for $1-\Phi(x)$ for large x so that we can accurately determine the asymptotic distribution of the maximum.

Theorem 3.1. *Let* $V = -2\log(1-\Phi(x)) - \log(2\pi)$*.*

$$\begin{array}{ll} \textit{For } x \geq 2, & V - \log V \leq x^2, \\ \textit{For } x \geq 1, & x^2 \leq V - \log V + \log V/V. \end{array}$$

Proof: The standard bounds from Abramowitz and Stegun(1972), p 932:
For $x \geq 1$, with $y = x^2$,

$$\begin{array}{rcl} \frac{\phi(x)}{x}(1-\frac{1}{x^2}) \leq & 1-\Phi(x) & \leq \frac{\phi(x)}{x}(1-\frac{1}{x^2}+\frac{3}{x^4}), \\ y + \log y - 2\log\left(1-\frac{1}{y}+\frac{3}{y^2}\right) \leq & V & \leq y + \log y - 2\log\left(1-\frac{1}{y}\right). \end{array} \tag{3.1}$$

We demonstrate the specified bounds by explicit calculation for moderate x, and by using the standard bounds for large x. For the lower bound, let $V - y - \log y = 2\log\frac{dV}{dy} = \Delta$ where, from (3.1), $\Delta \leq -2\log\left(1-\frac{1}{y})\right)$.

$$\begin{array}{rcl} y - V + \log V & = & \log(V/y) - \Delta \\ & \geq & \log(1+\log y/y) + 2\log\left(1-\frac{1}{y})\right) \\ & \geq & 0 \text{ for } y \geq 11. \end{array}$$

The lower bound is thus established for $y \geq 11$. The lower bound is exhibited in explicit calculation for $4 < y < 11$, so that the lower bound holds for $y \geq 4$ which is $x \geq 2$.

For the upper bound, using $V = y + \log y + \Delta$ and $\Delta \geq \frac{1}{y} - \frac{3}{y^2}$,

$$\begin{array}{rcl} y - V + \log V - \log V/V & = & \log(V/y) - \log V/V - \Delta \\ & = & \log(1+(\log y + \Delta)/y) - \log V/V - \Delta \\ & \leq & \log y/y - \log V/V - (1-1/y)\Delta \\ & \leq & (V-y)\log y/y^2 - (1-1/y)\Delta \\ & \leq & (\log y + \Delta)\log y/y^2 - (1-1/y)\Delta \\ & \leq & (\log y/y)^2 - (1-1/y-\log y/y^2)(\frac{1}{y}-\frac{3}{y^2}) \\ & \leq & 0 \text{ for } y > 5 \end{array}$$

The upper bound is thus established for $y \geq 5$. The upper bound is exhibited in explicit calculation for $1 < y < 5$, so that the upper bound holds for $x \geq 1$.

Theorem 3.2. *Let M_n be the maximum of n independent unit Gaussians. Let $N = \log(n^2/2\pi)$. For each n, there exists a Gumbel variable G, a monotone function of M_n, such that , for $M_n \geq 2$,*

$$(N + 2G) - \log(2G + N) \leq M_n^2 \leq V - \log V + \log V/V \tag{3.2}$$

where $V = N + 2G + 2\exp(-G)/n$.

Proof: Substitute the Gaussian probability bounds from lemma 1 into the probability bounds for $\Phi(M_n)$ given in Theorem 3.1.

We see from theorem 3.2, that as $n \to \infty$, $M_n^2 - N - \log N \to 2G$ in distribution. Note that the bounds apply only to tail probabilities $\Pr\{M_n \geq A\}$ for $A > 2$.

Theorem 3.3. *Let M_n denote the maximum of n unit gaussian random variables. Let $N = \log(n^2/2\pi)$. Then there is an exponential variable E with*

$$M_n^2 \leq \max\left(1, N + 2E - \log(N + 2E) + \log(N + 2E)/(N + 2E)\right).$$

Proof: From theorem 3.1, with $V = -\log(2\pi) - 2\log(1 - \Phi(M_n))$, we have, for $M_n \geq 1$, $M_n^2 \leq V - \log V + \log V/V$. From theorem 1, there exists a function $E(M_n)$ distributed exponentially, with $V \leq N + 2E$. Thus,

$$M_n^2 \leq \max\left(1, N + 2E - \log(N + 2E) + \log(N + 2E)/(N + 2E)\right).$$

Asymptotically, $M_n^2 - N - \log N \leq 2E$ as $n \to \infty$.

Theorem 3.4. *Let M_n be the maximum of n zero-mean gaussian variables $X_1, .., X_n$.*
Define

$$\begin{aligned} E_i &= \Pr(X_i | X_1, .., X_{i-1}), \\ R_i &= X_i - E_i, \\ \tau^2 &= \max_{1 \leq i \leq n} \tau_i^2 = \Pr E_i^2, \\ \sigma^2 &= \min_{1 \leq i \leq n} \sigma_i^2 = \Pr R_i^2, \\ N &= \log(n^2/2\pi), \\ L_\alpha &= -2\log(-\log\alpha). \end{aligned}$$

Then, for $N + L_\alpha \geq 6$,

$$\Pr\{M_n \geq \sigma(N + L_\alpha - \log(N + L_\alpha)^{\frac{1}{2}} + \tau\Phi^{-1}(\alpha)\} \geq 1 - 2\alpha.$$

Proof: We first show that for each real A and non-positive B,

$$\Pr\{M_n \geq A + B\} \geq \Pr\{\max_i R_i \geq A\} \min_i \Pr\{E_i \geq B\}.$$

Construct n disjoint events

$$H_i = \{R_i \geq A\} \prod_{j>i} \{R_j < A\}.$$

Note that $\sum_i H_i = \{\max_i R_i \geq A\}$. Since the terms $\{R_i \geq A\}$ are independent of all variables $X_j, j < i$, the individual terms in each H_i are l independent, and H_i is independent of $\{E_i \geq B\}$. Also, $\{M_n \geq A + B\} \geq \sum_i H_i\{E_i \geq B\}$, since at most one of the events $H_i\{E_i \geq B\}$ occurs, and if any occurs $M_n \geq A + B$. Then

$$\begin{aligned}
\Pr\{M_n \geq A + B\} &\geq \sum_i \Pr H_i \Pr\{E_i \geq B\} \\
&\geq \sum_i \Pr H_i \min_j \Pr\{E_j \geq B\} \\
&\geq \Pr\{\max_i R_i \geq A\} \min_i \Pr\{E_i \geq B\}.
\end{aligned}$$

For $A > 0$, $\{\max_i R_i \geq A\} \geq \{\sigma \max_i (R_i/\sigma_i) \geq A\}$.
Let $R_{(n)} = \sigma \max_i (R_i/\sigma_i)$. From theorem 4, for $A/\sigma \geq 2$,

$$\Pr\{R_{(n)} \geq A\} \geq \Pr\{(N + 2G - \log(N + 2G))^{\frac{1}{2}} \geq A/\sigma\}.$$

Since $\Pr\{2G \geq L_\alpha\} = 1 - \alpha$, for $N + L_\alpha \geq 6$ (which implies $A/\sigma \geq 2$),

$$\Pr\{R_{(n)} \geq \sigma(N + L_\alpha - \log(N + L_\alpha))^{\frac{1}{2}}\} \geq 1 - \alpha.$$

Also $\min_i \Pr\{E_i \geq B\} = 1 - \Phi(B/\tau)$, so

$$\min_i \Pr\{E_i \geq \tau\Phi^{-1}(\alpha)\} = 1 - \alpha.$$

Combining the two bounds gives, for $N + L_\alpha \geq 6$,

$$\Pr\{M_n \geq \sigma(N + L_\alpha - \log(N + L_\alpha)^{\frac{1}{2}} + \tau\Phi^{-1}(\alpha)\} \geq 1 - 2\alpha.$$

as asserted.

If the random variables $X_1, .., X_n$ have covariance matrix C, then

$$\sigma^2 = \min_i \frac{1}{C_{ii}^{-1}} \geq \min_x \frac{x'Cx}{x'x},$$

$$\tau^2 = \max_i \tau_i^2 \leq \max_i \Pr X_i^2 \leq \max_x \frac{x'Cx}{x'x}.$$

Thus the stated lower bound holds if the minimum eigenvalue of C is substituted for σ^2 and the maximum eigenvalue of C is substituted for τ^2.

In simulations, the lower bound is reasonably close to the actual distribution of M_n in the tail when the variables are i.i.d, but can be quite conservative when the minimum eigenvalue is small.

4. Application to stationary Gaussian processes

Following Wold[4], purely non-deterministic stationary Gaussian processes $X_i, \infty < i < \infty$, may be expressed in terms of the i.i.d innovations

$$Z_i = X_i - \Pr(X_i | X_{i-1}, X_{i-2}, ..) :$$

$$X_i = Z_i + \sum_{j=1}^{\infty} \psi_j Z_{i-j}.$$

Let X_i have variance 1. Theorem 5 now applies with $\sigma^2 = \Pr Z_0^2, \tau^2 = 1 - \sigma^2$. The inequality is dominated asymptotically by the first term, so that,

$$\text{for each } \epsilon > 0, \Pr\{M_n > \sqrt{2\sigma^2 \log n}(1-\epsilon)\} \to 0 \text{ as } n \to \infty.$$

The subsampled series $Y_i = X_{ik}$ for integer k has innovations $Z_i + \sum_{j=1}^{k-1} \psi_j Z_{i-j}$ with residual variances $1 - \sigma^2(\sum_{j=k}^{\infty} \psi_j^2)$, which may be chosen arbitrarily close to 1 by choosing k large enough. Since $M_n = \max_1^n X_i \geq \max_1^{n/k} Y_i$, the stated inequality for M_n holds with these residual variances.
Thus, for each $\epsilon > 0$, $M_n > \sqrt{2 \log n}(1-\epsilon)$ in probability as $n \to \infty$. (The order of magnitude based on Y is $\sqrt{2\log(n/k)}$, but the k washes out asymptotically.) From theorem 3.2, we have $M_n < \sqrt{2 \log n}(1+\epsilon)$ in probability. Thus $M_n / \sqrt{2 \log n} \to 1$ in probability, just as for an i.i.d. series. Detailed lower bounds for the tail probabilities of M_n may be obtained depending on the rate of convergence of the series $\sum_{j=1}^{\infty} \psi_j^2$.

References

[1] Abramowitz, M.and Stegun, I. A., eds. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York MR0208797
[2] Berman, S. I.(1964). Limit Theorems for the Maximum Term in Stationary Sequences *Ann. Math. Statist.* **35** 502-516. MR0161365
[3] Gumbel, E. J.(1941) The Return Period of Flood Flows *Ann. Math. Statist.* **12** 163-190. MR0004457
[4] Resnick, S. I.(1972) Stability of Maxima of Random Variables Defined on a Markov Chain *Adv. Appl. Prob.* **4** 284-295. MR0336820
[5] Wold, H. A.(1938) *Study in the Analysis of Stationary Time Series.* Almqvist and Wiksel, Uppsala