

Bounds on the Expectation of the Maximum of Samples from a Gaussian

Gautam Kamath

In this document, we will provide bounds on the expected maximum of n samples from Gaussian distribution. Let $Y = \max_{1 \leq i \leq n} X_i$, where X_i are iid random variables, distributed as $\mathcal{N}(0, \sigma^2)$. We will show that

$$\sigma \frac{1}{\sqrt{\pi \log 2}} \sqrt{\log n} \leq E[Y] \leq \sigma \sqrt{2} \sqrt{\log n}$$

First, we show $E[Y] \leq \sigma \sqrt{2} \sqrt{\log n}$.

$$\begin{aligned} \exp\{tE[Z]\} &\leq E[\exp\{tZ\}] \\ &= E[\max \exp\{tX_i\}] \\ &\leq \sum_{i=1}^n E[\exp\{tX_i\}] \\ &= n \exp\{t^2 \sigma^2 / 2\} \end{aligned}$$

The first inequality is Jensen's inequality, the second is the union bound, and the final equality follows from the definition of the moment generating function.

Taking the logarithm of both sides of this inequality, we get

$$E[Z] \leq \frac{\log n}{t} + \frac{t\sigma^2}{2}$$

This can be minimized by setting $t = \frac{\sqrt{2 \log n}}{\sigma}$, which gives us the desired result,

$$E[Z] \leq \sigma \sqrt{2} \sqrt{\log n}$$

Next, we show the more difficult direction - that $E[Y] \geq \sigma C \sqrt{\log n}$, where $C = \frac{1}{\sqrt{\pi \log 2}}$. If one of our n samples is greater than some value, then the maximum is at least that value. Thus, we lower bound the expectation of Y by showing that we expect at least one of the X_i to be greater than $\sigma C \sqrt{\log n}$.

Let $k = \frac{2C^2}{\pi}$.

$$\begin{aligned}
E[|\{i : X_i \geq \sigma C \sqrt{\log n}\}|] &= nP(X_i \geq \sigma C \sqrt{\log n}) \\
&= \frac{n}{2}P(|X_i| \geq \sigma C \sqrt{\log n}) \\
&= \frac{n}{2} \left(1 - \operatorname{erf} \left(\frac{C}{\sqrt{2}} \sqrt{\log n} \right) \right) \\
&\geq \frac{n}{2} \left(1 - \sqrt{1 - \exp(-k \log n)} \right) \\
&= \frac{n}{2} \left(1 - \sqrt{1 - n^{-k}} \right)
\end{aligned}$$

Where the second equality follows by symmetry, the third equality is based on the CDF of the half-normal distribution, and the inequality is from the bound on the error function, $\operatorname{erf}(x) \leq \sqrt{1 - \exp(-\frac{4}{\pi}x^2)}$. We require this value to be at least 1:

$$\begin{aligned}
\frac{n}{2} \left(1 - \sqrt{1 - n^{-k}} \right) &\geq 1 && \Leftrightarrow \\
1 - \frac{2}{n} &\geq \sqrt{1 - n^{-k}} && \Leftrightarrow \\
1 - \frac{4}{n} + \frac{4}{n^2} &\geq 1 - \frac{1}{n^k} && \Leftrightarrow \\
n^{2-k} &\geq 4n - 4 && \Leftrightarrow \\
(2 - k) \log n &\geq \log(4n - 4) && \Leftrightarrow \\
2 - \frac{\log(4n - 4)}{\log n} &\geq k
\end{aligned}$$

Substituting in the value $C = \frac{1}{\sqrt{\pi \log 2}}$ tell us this inequality holds for $n \geq 5.27$. Since we are only concerned with integer values of n , this inequality holds for $n \geq 6$. We can confirm the inequality for $1 \leq n \leq 5$ by comparing $\sigma C \sqrt{\log n}$ with the explicit equations for $E[Y]$. Note that this inequality is tight for $n = 2$. Therefore, we have $E[Y] \geq \sigma \frac{1}{\sqrt{\pi \log 2}} \sqrt{\log n}$.