

## RESEARCH ARTICLE

## Using the sample maximum to estimate the parameters of the underlying distribution

Alex Capaldi, Tiffany N. Kolba \*

Department of Mathematics and Statistics, Valparaiso University, Valparaiso, IN, United States of America

\* [tiffany.kolba@valpo.edu](mailto:tiffany.kolba@valpo.edu)

## OPEN ACCESS

**Citation:** Capaldi A, Kolba TN (2019) Using the sample maximum to estimate the parameters of the underlying distribution. PLoS ONE 14(4): e0215529. <https://doi.org/10.1371/journal.pone.0215529>

**Editor:** Eugene Demidenko, Dartmouth College Geisel School of Medicine, UNITED STATES

**Received:** October 24, 2018

**Accepted:** April 3, 2019

**Published:** April 25, 2019

**Copyright:** © 2019 Capaldi, Kolba. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript, Supporting Information files, and on the OSF data repository: <https://osf.io/zg4ka/>.

**Funding:** A.C. was supported by the National Science Foundation under grant IOS-1645508 (<https://www.nsf.gov/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

We propose novel estimators for the parameters of an exponential distribution and a normal distribution when the only known information is a sample of sample maxima; i.e., the known information consists of a sample of  $m$  values, each of which is the maximum of a sample of  $n$  independent random variables drawn from the underlying exponential or normal distribution. We analyze the accuracy and precision of the estimators using extreme value theory, as well as through simulations of the sampling distributions. For the exponential distribution, the estimator of the mean is unbiased and its variance decreases as either  $m$  or  $n$  increases. Likewise, for the normal distribution, we show that the estimator of the mean has negligible bias and the estimator of the variance is unbiased. While the variance of the estimators for the normal distribution decreases as  $m$ , the number of sample maxima, increases, the variance increases as  $n$ , the sample size over which the maximum is computed, increases. We apply our method to estimate the mean length of pollen tubes in the flowering plant *Arabidopsis thaliana*, where the known biological information fits our context of a sample of sample maxima.

## 1 Introduction

Consider the scenario where one has obtained data where each observation is the maximum value of  $n$  independent, identically distributed random variables drawn from either an exponential distribution or a normal distribution with unknown parameters. That is,  $X_{ij} \stackrel{\text{iid}}{\sim} \text{Exp}(\beta)$  or  $X_{ij} \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$  for  $i = 1, \dots, n$  and known data is drawn from  $Y_j = \max\{X_{ij}\}_{i=1}^n$  for  $j = 1, \dots, m$ . Here we present a process to estimate the mean  $\beta$  of the underlying exponential distribution or the mean  $\mu$  and variance  $\sigma^2$  of the underlying normal distribution from only the set of  $Y_j$ 's.

Much previous research has been conducted in the field of extreme value theory on the distribution of the maximum from a sample of  $n$  independent, identically distributed random variables. In particular, the Fisher-Tippett-Gnedenko theorem states that the distribution of the sample maximum, after proper rescaling, can only converge to one of three types of distributions: the Gumbel distribution, the Fréchet distribution, or the Weibull distribution [1]. When the original underlying distribution is exponential or normal, the limiting distribution of the rescaled sample maximum is the Gumbel distribution. Extreme value theory has been

applied in many applications, such as estimating the probability of an extreme flood, severe adverse side effect of a drug, maximum environmental load on a structure, or large insurance loss [2]. In these applications, the underlying distribution and its parameters are typically known and the focus is on estimating the probability distribution of the sample maximum. The focus of our scenario is novel in that we are using a known sample of sample maxima to estimate unknown parameters of the underlying distribution.

Standard techniques for estimating unknown parameters, such as method of moments or maximum likelihood estimation, typically assume that the known information consists of a sample of observations drawn directly from the underlying population distribution. However, in our scenario under consideration, the direct observations are unknown. Rather, we only know the maximum value of each sample of direct observations. The estimators we propose in the following sections are the first, to our knowledge, for estimating unknown population parameters when the only known information is a sample of sample maxima.

## 2 Estimator for exponential distribution

We begin by considering the case where the underlying distribution is exponential with unknown mean  $\beta$ . In Theorem 1 below, we propose an estimator for  $\beta$  and compute its expected value and variance.

**Theorem 1.** Let  $X_{ij} \stackrel{iid}{\sim} Exp(\beta)$  for  $i = 1, \dots, n$  and  $Y_j = \max\{X_{ij}\}_{i=1}^n$  for  $j = 1, \dots, m$ . Set

$$\hat{\beta} = \frac{\bar{Y}}{H_n} = \frac{1}{mH_n} \sum_{j=1}^m Y_j. \tag{1}$$

Then

$$E(\hat{\beta}) = \beta, \quad var(\hat{\beta}) = \frac{\beta^2 G_n}{mH_n^2} \tag{2}$$

where

$$H_n = \sum_{i=1}^n \frac{1}{i}, \quad G_n = \sum_{i=1}^n \frac{1}{i^2}.$$

*Proof.* From the formula for  $\hat{\beta}$  given in Eq (1), it directly follows that

$$E(\hat{\beta}) = \frac{E(Y_j)}{H_n}, \quad var(\hat{\beta}) = \frac{var(Y_j)}{mH_n^2}.$$

Hence, it only remains to compute the expected value and variance of the maximum of a single sample of  $n$  independent  $Exp(\beta)$  random variables. Let  $X_{(i)}$  denote the  $i$ th smallest observation from such a sample. Then  $Y_j = X_{(n)}$  can be decomposed as the following telescoping sum:

$$Y_j = X_{(n)} = X_{(1)} + (X_{(2)} - X_{(1)}) + \dots + (X_{(n)} - X_{(n-1)}).$$

Due to the memoryless property of the exponential distribution,  $X_{(2)} - X_{(1)}$  is independent of  $X_{(1)}$ . Moreover, while  $X_{(1)}$  is the minimum of  $n$  independent  $Exp(\beta)$  random variables,  $X_{(2)} - X_{(1)}$  can be viewed as the minimum of a sample of  $n - 1$  independent  $Exp(\beta)$  random variables. Likewise, all of the terms in the telescoping sum for  $Y_j = X_{(n)}$  are independent with  $X_{(i+1)} - X_{(i)}$  equal in distribution to the minimum of a sample of  $n - i$  independent  $Exp(\beta)$

random variables, which in turn is equal in distribution to  $Exp(\beta/(n - i))$ . Thus,

$$E(Y_j) = E(X_{(n)}) = \frac{\beta}{n} + \frac{\beta}{n-1} + \dots + \frac{\beta}{2} + \beta = \beta \sum_{i=1}^n \frac{1}{i} = \beta H_n$$

and

$$\text{var}(Y_j) = \text{var}(X_{(n)}) = \frac{\beta^2}{n^2} + \frac{\beta^2}{(n-1)^2} + \dots + \frac{\beta^2}{2^2} + \beta^2 = \beta^2 \sum_{i=1}^n \frac{1}{i^2} = \beta^2 G_n.$$

Substituting these expressions for  $E(Y_j)$  and  $\text{var}(Y_j)$  into the equations for the expected value and variance of  $\hat{\beta}$  produces the formulas given in Eq (2).

As a consequence of Theorem 1, we have shown that  $\hat{\beta}$  is an unbiased estimator for  $\beta$  and that its variance decreases at a rate proportional to  $\frac{1}{m}$ . Since  $G_n \rightarrow \frac{\pi^2}{6}$  and  $H_n \rightarrow \infty$  at rate  $\log n$  as  $n \rightarrow \infty$ , the variance of  $\hat{\beta}$  decreases at a rate proportional to  $\frac{1}{(\log n)^2}$ . Thus, the precision of the estimator  $\hat{\beta}$  can be improved more rapidly by increasing  $m$ , the number of sample maxima, compared to increasing  $n$ , the sample size over which the maximum is computed.

### 3 Estimators for normal distribution

We now consider the case where the underlying distribution is normal with unknown mean  $\mu$  and unknown variance  $\sigma^2$ . In Theorem 2 below, we propose estimators for  $\mu$  and  $\sigma^2$  and analyze their expected value, while in Theorem 3 we analyze the variance of the estimators.

**Theorem 2.** Let  $X_{ij} \stackrel{iid}{\sim} N(\mu, \sigma^2)$  for  $i = 1, \dots, n$  and  $Y_j = \max\{X_{ij}\}_{i=1}^n$  for  $j = 1, \dots, m$ . Let  $\bar{Y}$  and  $S_Y^2$  denote the sample mean and sample variance, respectively, of the  $Y_j$ 's. Set

$$\hat{\mu} = \bar{Y} - \frac{k_n}{\sqrt{c_n}} S_Y, \quad \hat{\sigma}^2 = \frac{S_Y^2}{c_n} \tag{3}$$

where  $k_n$  denotes the mean and  $c_n$  denotes the variance of the maximum of  $n$  independent, identically distributed  $N(0, 1)$  random variables. Then  $E(\hat{\sigma}^2) = \sigma^2$ , while  $E(\hat{\mu}) > \mu$  with  $E(\hat{\mu}) \rightarrow \mu$  as  $m \rightarrow \infty$ .

*Proof.* The cumulative distribution function of  $Y_j$  is given by

$$\begin{aligned} F_{Y_j}(y) &= P(Y_j \leq y) = [P(X_{ij} \leq y)]^n \\ &= \left[ P\left(\frac{X_{ij} - \mu}{\sigma} \leq \frac{y - \mu}{\sigma}\right) \right]^n = \left[ \Phi\left(\frac{y - \mu}{\sigma}\right) \right]^n. \end{aligned}$$

Differentiating, the probability density function of  $Y_j$  is

$$f_{Y_j}(y) = n \left[ \Phi\left(\frac{y - \mu}{\sigma}\right) \right]^{n-1} \phi\left(\frac{y - \mu}{\sigma}\right) \frac{1}{\sigma},$$

where  $\Phi(z)$  denotes the cumulative distribution function and  $\phi(z)$  denotes the probability density function of a  $N(0, 1)$  random variable. The expected value of  $Y_j$  can then be calculated as

$$\begin{aligned} E(Y_j) &= \int_{-\infty}^{\infty} y f_{Y_j}(y) dy = \int_{-\infty}^{\infty} y n \left[ \Phi\left(\frac{y - \mu}{\sigma}\right) \right]^{n-1} \phi\left(\frac{y - \mu}{\sigma}\right) \frac{1}{\sigma} dy \\ &= \mu + \sigma \int_{-\infty}^{\infty} z n [\Phi(z)]^{n-1} \phi(z) dz = \mu + \sigma k_n. \end{aligned}$$

We also obtain that the variance of  $Y_j$  is

$$\begin{aligned} \text{var}(Y_j) &= \int_{-\infty}^{\infty} (y - E(Y_j))^2 f_{Y_j}(y) dy \\ &= \int_{-\infty}^{\infty} (y - \mu - \sigma k_n)^2 n \left[ \Phi\left(\frac{y - \mu}{\sigma}\right) \right]^{n-1} \phi\left(\frac{y - \mu}{\sigma}\right) \frac{1}{\sigma} dy \\ &= \sigma^2 \int_{-\infty}^{\infty} (z - k_n)^2 n [\Phi(z)]^{n-1} \phi(z) dz = \sigma^2 c_n. \end{aligned}$$

We can now use the equations for  $E(Y_j)$  and  $\text{var}(Y_j)$  to compute the expected value of the estimators. In particular, the expected value of the estimator of the variance is

$$E(\hat{\sigma}^2) = \frac{E(S_Y^2)}{c_n} = \frac{\text{var}(Y_j)}{c_n} = \frac{\sigma^2 c_n}{c_n} = \sigma^2,$$

while the estimator of the mean is

$$E(\hat{\mu}) = E(Y_j) - \frac{k_n}{\sqrt{c_n}} E(S_Y) > (\mu + \sigma k_n) - \frac{k_n}{\sqrt{c_n}} (\sigma \sqrt{c_n}) = \mu.$$

Although Jensen’s inequality implies that  $\hat{\mu}$  has positive bias since  $E(S_Y) < \sqrt{E(S_Y^2)} = \sqrt{\text{var}(Y_j)}$ , the bias of the sample standard deviation goes to zero as the sample size increases. Hence,  $E(\hat{\mu}) \rightarrow \mu$  as  $m \rightarrow \infty$ .

Note that the constants  $k_n$  and  $c_n$  that appear in the formulas for the estimators depend upon only the sample size  $n$ . Exact integral expressions exist for  $k_n$  and  $c_n$  and are given in the proof of Theorem 2, but the integrals cannot be evaluated in closed form. However, the constants can be approximated either analytically or numerically.

In [3], Cramér showed that  $b_n(Z_{(n)} - b_n)$  converges in distribution to the standard Gumbel distribution, where  $Z_{(n)}$  is the maximum of a sample of  $n$  independent, identically distributed  $N(0, 1)$  random variables and

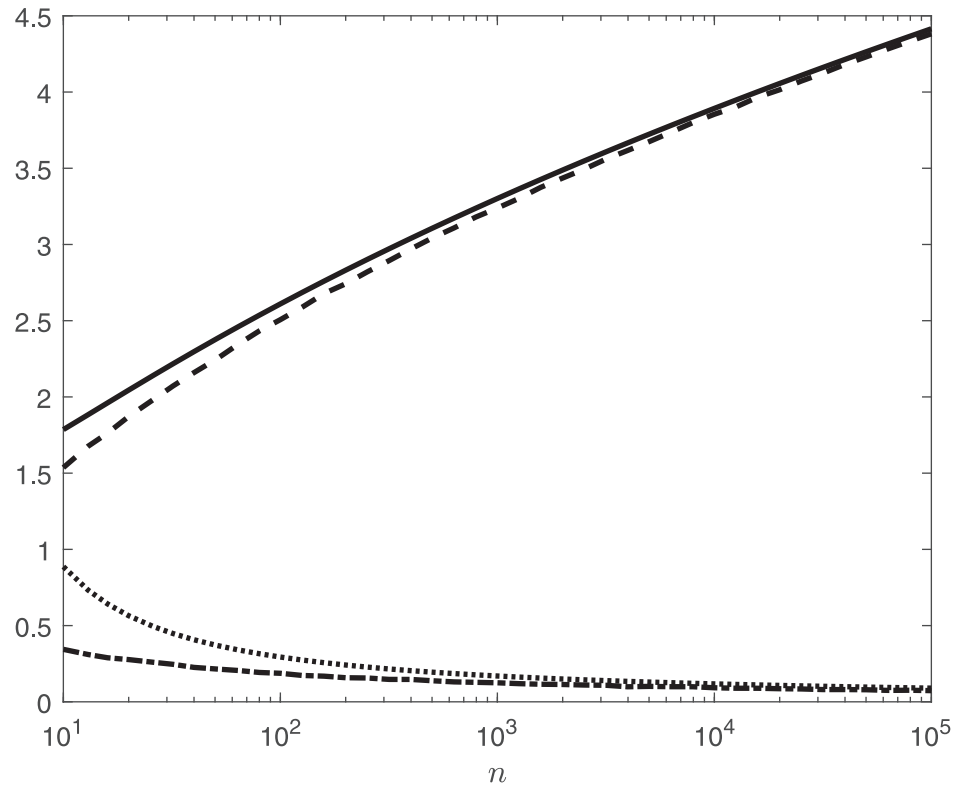
$$b_n = (2 \log n)^{\frac{1}{2}} - \frac{\frac{1}{2} \log(4\pi \log n)}{(2 \log n)^{\frac{1}{2}}}.$$

Since the standard Gumbel distribution has mean equal to  $\gamma \approx 0.5772$ , the Euler-Mascheroni constant, and variance equal to  $\frac{\pi^2}{6}$ , we can use these values to approximate

$$k_n \approx \frac{\gamma}{b_n} + b_n, \quad c_n \approx \frac{\pi^2}{6b_n^2}. \tag{4}$$

Fig 1 displays the values of the analytic approximations for  $k_n$  and  $c_n$  given in Eq (4) for  $n$  ranging from 10 to 100,000, along with corresponding numerical approximations, plotted on a semi-log scale. The numerical approximations for  $k_n$  and  $c_n$  were computed from 10,000 realizations of a simulation of the maximum.

While Theorem 2 showed that  $\hat{\mu}$  is positively biased with the bias approaching zero as  $m \rightarrow \infty$ , the estimation bias is fairly minimal even for relatively small values of  $m$ . Fig 2 displays the sampling distributions of the estimators for  $\mu$  and  $\sigma^2$  for varying values of  $m$  and  $n$ . In all the simulations, the sampling distribution of  $\hat{\mu}$  is fairly centered around the true value of  $\mu = 0$ . Setting the true value of  $\mu$  to a nonzero value simply shifts the sampling distribution of  $\hat{\mu}$  and has no effect on  $\hat{\sigma}^2$ . We also observe from Fig 2 that the variability of the sampling distributions of the estimators decreases as  $m$ , the number of sample maxima, increases, but



**Fig 1.** From the top curve to the bottom, the plot displays the values of the analytic approximation for  $k_n$  (solid) and the numerical approximation for  $k_n$  (dashed), along with the analytic approximation for  $c_n$  (dotted) and the numerical approximation for  $c_n$  (dot-dashed).

<https://doi.org/10.1371/journal.pone.0215529.g001>

increases as  $n$ , the sample size over which the maximum is computed, increases. We derive an analytical justification for this behavior in Theorem 3 below.

**Theorem 3.** Let  $X_{ij} \stackrel{iid}{\sim} N(\mu, \sigma^2)$  for  $i = 1, \dots, n$  and  $Y_j = \max\{X_{ij}\}_{i=1}^n$  for  $j = 1, \dots, m$ . Let  $\bar{Y}$  and  $S_Y^2$  denote the sample mean and sample variance, respectively, of the  $Y_j$ 's. Set

$$\hat{\mu} = \bar{Y} - \frac{k_n}{\sqrt{c_n}} S_Y, \quad \hat{\sigma}^2 = \frac{S_Y^2}{c_n}$$

where  $k_n$  denotes the mean and  $c_n$  denotes the variance of the maximum of  $n$  independent, identically distributed  $N(0, 1)$  random variables. Then

$$\text{var}(\hat{\sigma}^2) = \sigma^4 \left( \frac{2}{m-1} + \frac{\kappa}{m} \right)$$

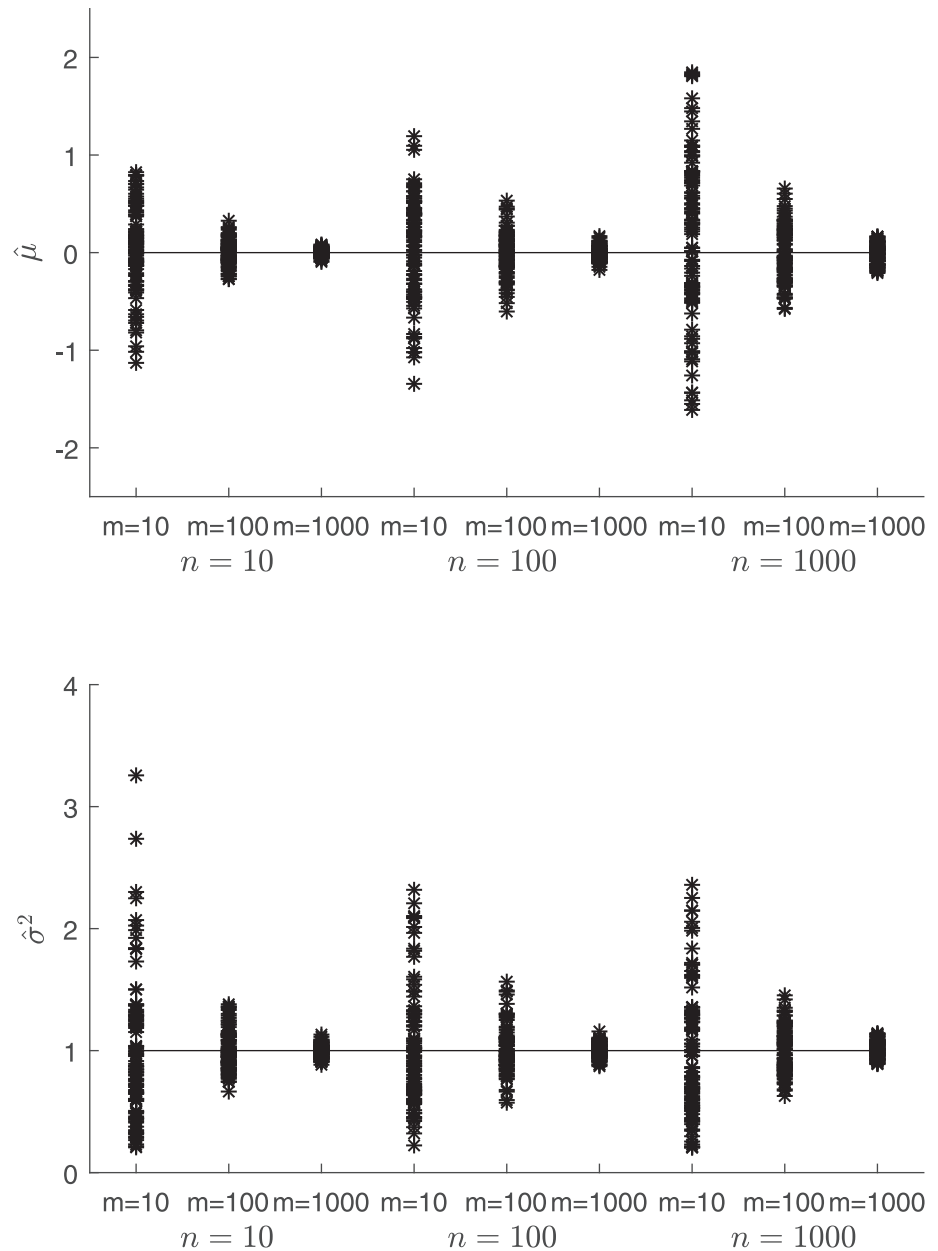
and

$$\text{var}(\hat{\mu}) \approx \frac{\sigma^2 c_n}{m} + \frac{\sigma^2 k_n^2}{4} \left( \frac{2}{m-1} + \frac{\kappa}{m} \right) - \frac{\sigma^2 k_n \sqrt{c_n} \gamma_1}{m}$$

where  $\gamma_1$  is the skewness and  $\kappa$  is the excess kurtosis of the distribution of  $Y_j$ .

*Proof.* The variance of  $S_Y^2$  can be computed as

$$\text{var}(S_Y^2) = (\text{var}(Y_j))^2 \left( \frac{2}{m-1} + \frac{\kappa}{m} \right),$$



**Fig 2. Estimates of  $\hat{\mu}$  (top) and  $\hat{\sigma}^2$  (bottom) from 100 realizations with  $n, m = 10, 100,$  and  $1000$ . The horizontal lines indicate the true values of  $\mu$  and  $\sigma^2$ .**

<https://doi.org/10.1371/journal.pone.0215529.g002>

where  $\kappa$  is the excess kurtosis of the distribution of  $Y_j$  [4]. Since we computed that  $\text{var}(Y_j) = \sigma^2 c_n$  in Theorem 2, we obtain the desired result

$$\text{var}(\hat{\sigma}^2) = \text{var}\left(\frac{S_Y^2}{c_n}\right) = \sigma^4 \left(\frac{2}{m-1} + \frac{\kappa}{m}\right)$$

for the variance of the variance estimator. Now for the variance of the mean estimator, we

compute that

$$\text{var}(\hat{\mu}) = \text{var}\left(\bar{Y} - \frac{k_n}{\sqrt{c_n}} S_Y\right) = \text{var}(\bar{Y}) + \frac{k_n^2}{c_n} \text{var}(S_Y) - 2 \frac{k_n}{\sqrt{c_n}} \text{cov}(\bar{Y}, S_Y).$$

To simplify the above expression, we use the fact that  $\text{var}(\bar{Y}) = \frac{\text{var}(Y_j)}{m}$ , along with the approximations

$$\text{var}(S_Y) \approx \frac{\text{var}(S_Y^2)}{4\text{var}(Y_j)}, \quad \text{cov}(\bar{Y}, S_Y) \approx \frac{\text{cov}(\bar{Y}, S_Y^2)}{2\sqrt{\text{var}(Y_j)}} = \frac{\text{var}(Y_j)\gamma_1}{2m},$$

where  $\gamma_1$  is the skewness of the distribution of  $Y_j$  [5]. Using these approximations, we obtain the desired result

$$\text{var}(\hat{\mu}) \approx \frac{\sigma^2 c_n}{m} + \frac{\sigma^2 k_n^2}{4} \left(\frac{2}{m-1} + \frac{\kappa}{m}\right) - \frac{\sigma^2 k_n \sqrt{c_n} \gamma_1}{m}$$

for the variance of the mean estimator.

As  $n$  increases, the value of  $\kappa$  increases monotonically from 0, the excess kurtosis of the normal distribution, to 12/5, the excess kurtosis of the Gumbel distribution [1]. Thus, from Theorem 3, we observe that the variance of the sampling distribution of  $\hat{\sigma}^2$  increases proportionally to  $\sigma^4$  as  $\sigma$  increases and decreases proportionally to  $1/m$  as  $m$  increases, but only slightly increases as  $n$  increases.

As with the excess kurtosis, the skewness of the distribution of  $Y_j$  increases monotonically from the value for the normal distribution, i.e.,  $\gamma_1 = 0$ , to the value for the standard Gumbel distribution, i.e.,  $\gamma_1 \approx 1.13955$ , as  $n$  increases [6]. Since the constant  $c_n$  decreases towards 0 while the constant  $k_n$  increases towards infinity, the dominant term in the variance of  $\hat{\mu}$  increases proportionally to  $k_n^2$  as  $n$  increases. We also observe from Theorem 3 that the variance of the sampling distribution of  $\hat{\mu}$  increases proportionally to  $\sigma^2$  as  $\sigma$  increases and decreases proportionally to  $1/m$  as  $m$  increases. These relationships explain the behavior of the sampling distributions displayed in Fig 2.

## 4 Biological application

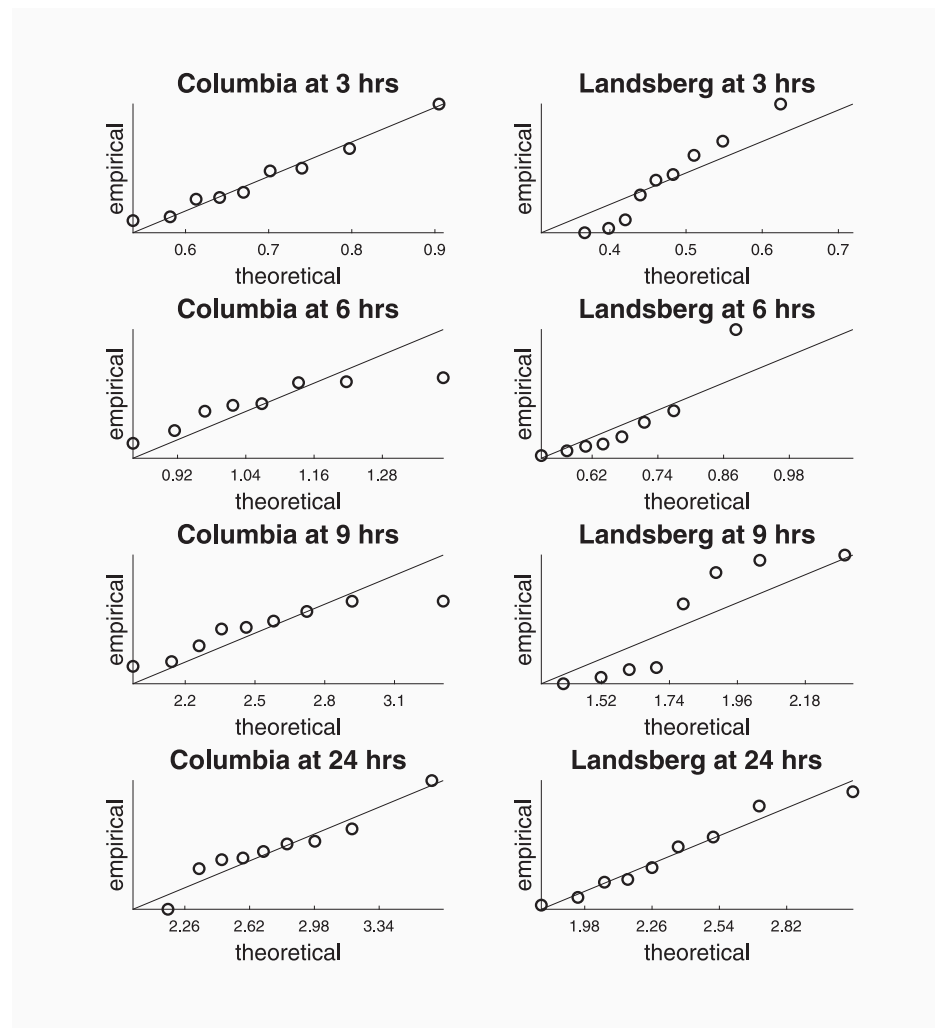
During fertilization in flowering plants, once pollen land on the stigma, the pollen will grow tubes that travel down through a transmitting tract from the stigma toward an ovule. Pollen compete against each other in a race towards the limited number of ovules to determine which pollen will father the seeds. The mean length of the population of pollen tubes at various time points is of interest to plant biologists, yet, to date, there are only measures of the lengths of the *longest* pollen tubes in such competitions [7]. Since the pollen tube lengths must have a positive value, it is reasonable to assume that the lengths follow an exponential distribution. Hence, our method described in Section 2 will allow the mean pollen tube length to be estimated given the structure of the experimental data.

In [7], Swanson et al. measured the longest pollen tube lengths at four time points for two accessions (i.e., specific geographical populations) of *Arabidopsis thaliana* in a laboratory setting. For both the Columbia and Landsberg accessions, either  $m = 8$  or  $m = 9$  individual plants were used for each time point. The average number of pollen tubes within each plant was  $n = 933$  for the Columbia accession and  $n = 727$  for the Landsberg accession. Table 1 reports the sample mean of the longest pollen tube from the  $m$  plants for each accession after 3, 6, 9, and 24 hours. Using these sample means of the longest lengths and Eq (2) from Theorem 1, we

**Table 1.** The sample mean,  $\bar{y}$ , of the longest pollen tube from  $m$  plants, where each plant contained an average of  $n$  pollen tubes, along with the corresponding estimated mean pollen length,  $\hat{\beta}$ , and its standard error, for two accessions of *Arabidopsis thaliana* at varying time points.

Accession	time (hrs)	$m$	$\bar{y}$ (mm)	$\hat{\beta}$ (mm)	$SE_{\hat{\beta}}$ (mm)
Columbia ( $n = 933$ )	3	9	0.690	0.093	0.005
	6	8	1.069	0.144	0.009
	9	9	2.538	0.342	0.020
	24	9	2.778	0.375	0.022
Landsberg ( $n = 727$ )	3	9	0.474	0.066	0.004
	6	8	0.676	0.094	0.006
	9	8	1.795	0.251	0.016
	24	9	2.325	0.324	0.019

<https://doi.org/10.1371/journal.pone.0215529.t001>



**Fig 3.** Q-Q plots of the distribution of the maximum pollen tube length from laboratory experiments versus the distribution of the maximum from an exponential distribution.

<https://doi.org/10.1371/journal.pone.0215529.g003>



then estimated the overall mean length for each accession at each time point. The resulting estimates and their standard errors are listed in [Table 1](#).

To further evaluate the validity of the assumption that the population of pollen tube lengths is exponentially distributed, we produced Q-Q plots of the distribution of the maximum pollen tube length from the laboratory experiments performed by Swanson et al. versus the distribution of the maximum from an exponential distribution. The theoretical distribution of the maximum from an exponential distribution was simulated using 10,000 realizations of  $Y_j = \max\{X_{ij}\}_{i=1}^n$  where  $X_{ij} \stackrel{\text{iid}}{\sim} \text{Exp}(\hat{\beta})$  for  $i = 1, \dots, n$ , using the values of  $n$  and  $\hat{\beta}$  that are listed in [Table 1](#). The Q-Q plots, displayed in [Fig 3](#), show a roughly linear relationship, supporting the assumption that the underlying distribution of pollen tube lengths is exponential. Moreover, we performed a Kolmogorov-Smirnov test of the equality of the empirical and theoretic distributions for each accession of *Arabidopsis thaliana* at each time point. The smallest resulting p-value was 0.52 (corresponding to the Landsberg accession at 9 hours), further indicating that there is no evidence that the distribution of pollen tube lengths differs significantly from an exponential distribution.

## Acknowledgments

This work was supported by the National Science Foundation under grant IOS-1645508. We would also like to thank Swanson et al. for providing the data for the biological example.

## Author Contributions

**Conceptualization:** Alex Capaldi.

**Formal analysis:** Tiffany N. Kolba.

**Funding acquisition:** Alex Capaldi.

**Methodology:** Tiffany N. Kolba.

**Software:** Alex Capaldi.

**Visualization:** Alex Capaldi.

**Writing – original draft:** Tiffany N. Kolba.

## References

1. David HA, Nagaraja HN. Order statistics. Wiley Online Library; 1970.
2. Castillo E, Hadi AS, Balakrishnan N, Sarabia JM. Extreme value and related models with applications in engineering and science. Wiley Hoboken, NJ; 2005.
3. Cramér H. Mathematical methods of statistics. Princeton University Press; 1946.
4. Mood AM, Graybill FA, Boes DC. Introduction to the theory of statistics. McGraw-Hill; 1974.
5. Rao CR. Linear statistical inference and its applications, vol 2. Wiley New York; 1973.
6. Gumbel EJ. Statistics of extremes. Courier Corporation; 2012.
7. Swanson RJ, Hammond AT, Carlson AL, Gong H, Donovan TK. Pollen performance traits reveal prezygotic nonrandom mating and interference competition in *Arabidopsis thaliana*. American Journal of Botany. 2016 Feb; 103(3):498–513. <https://doi.org/10.3732/ajb.1500172> PMID: 26928008