Some Transformed Distributions

by

Hassan Abuhassan

Master of Science in Mathematics, The Jordanian University, 1991

A Dissertation
Submitted in Partial Fulfillment of the Requirements for the
Doctor of Philosophy Degree

Department of Mathematics
in the Graduate School
Southern Illinois University Carbondale
October, 2007

DISSERTATION APPROVAL

SOME TRANSFORMED DISTRIBUTIONS By

Hassan Abuhassan

A Dissertation Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in the field of Mathematics

Approved by:

David Olive, Chair

Bhaskar Bhattacharya

Edward Neuman

Salah-Eldin Mohammed

Todd Headrick

Graduate School
Southern Illinois University Carbondale
October 18, 2007

AN ABSTRACT OF THE DISSERTATION OF

Hassan Abuhassan, for the Doctor of Philosophy degree in MATHEMATICS, presented on October 00, 2007, at Southern Illinois University Carbondale.

TITLE: Some Transformed Distributions

MAJOR PROFESSOR: Dr. David. Olive

There are several useful asymmetric location-scale families that are one parameter exponential families when the location parameter is known. In this case inference is simple and the maximum likelihood estimator (MLE) and uniformly minimum variance unbiased estimator (UMVUE) are important point estimators. The Burr, largest extreme value, Pareto, power, Rayleigh, smallest extreme value, truncated extreme value, and Weibull distributions are obtained by transforming the exponential distribution. By applying the same transformation to the half normal distribution, eight new competitors for these distributions are obtained.

Inference for some of these transformed distributions is simple using inference for the original distributions and the invariance principle. Pewsey [15] studied the half normal distribution $HN(\mu, \sigma^2)$ and gave confidence intervals for the parameters, we give a better confidence interval for μ.

We also studied the Pareto, Rayleigh, and Weibull distributions and give confidence intervals for the parameters. In the case of the Pareto distribution $Pareto(\sigma, \lambda)$, the obtained confidence intervals for σ and λ seem to be new, while in the case of the Weibull distribution $Weibull(\phi, \lambda)$ our contribution was that we used robust estimators for ϕ and λ which are used in the iteration procedures to find the MLEs.

DEDICATION

To my father who taught me that the sky is the limit

To my wife who standed side by side with me

ACKNOWLEDGMENTS

I would like to thank Dr. David Olive for his invaluable assistance and insights leading to the writing of this paper. My sincere thanks also goes to Dr. Bhaskar Bhattacharya, Dr. Edward Neuman, Dr. Salah-Eldin Mohammed, and Dr. Todd Headrick, for serving as my committee members.

I would like also to thank the many people at the Department of Mathematics who contributed to my work in so many ways. Special thanks to Diane Fritcher for your endless assistance, and Linda Gibson for your patient help on my thesis-writing.

At last, I would like to thank all my friends and family, particularly my father, my wife, and my kids for their patience and support, my sisters and their families for their help and hospitality.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 INTRODUCTION

There are several useful asymmetric location-scale families that are one parameter exponential families when the location parameter is known. In this case inference is simple and the maximum likelihood estimator (MLE) and uniformly minimum variance unbiased estimator (UMVUE) are important point estimators. The Burr, largest extreme value, Pareto, power, Rayleigh, smallest extreme value, truncated extreme value, and Weibull distributions are obtained by transforming the exponential distribution. By applying the same transformation to the half normal distribution, eight new competitors for these distributions are obtained.

Inference for some of these transformed distributions is simple using inference for the original distributions and the invariance principle. Pewsey [15] studied the half normal distribution $HN(\mu, \sigma^2)$ and gave confidence intervals for the parameters, we give a better confidence interval for μ.

We also studied the Pareto, Rayleigh, and Weibull distributions and give confidence intervals for the parameters. In the case of the Pareto distribution $Pareto(\sigma, \lambda)$ which is studied extensively by Arnold (1983) [1], the obtained confidence intervals for σ and λ seem to be new, while in the case of the Weibull distribution $Weibull(\phi, \lambda)$ our contribution was that we used robust estimators for ϕ and λ which are used in the iteration procedures to find the MLEs.

Definition 1. The *population median* is any value MED(Y) such that $P(Y \leq \text{MED(Y)}) \geq 0.5$, and $P(Y \geq \text{MED(Y)}) \geq 0.5$.

Definition 2. The *population median absolute deviation* is MAD(Y) = MED(|Y − MED(Y)|).

Definition 3. A family of pdf's (probability density functions) or pmf's (probability mass functions) $f(x, \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta$ is an exponential family if

$$f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta}) \exp[\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)] \tag{1.1}$$

for $x \in \mathcal{X}$ where $c(\boldsymbol{\theta}) \geq 0$, $h(x) \geq 0$ does not depend on $\boldsymbol{\theta}$ and $t_i(x) : \mathcal{X} \to R$ does not depend on $\boldsymbol{\theta}$. The family is a k-parameter exponential family if k is the smallest integer where the above equation holds. If $k = 1$ and the exponential family is regular then it is called one parameter regular exponential family denoted by a **1P–REF**.

Suppose that $Y = t(W)$ and $W = t^{-1}(Y)$ where W has a pdf with parameters $\boldsymbol{\theta}$, the transformation t does not depend on any unknown parameters, and the pdf of Y is

$$f_Y(y) = f_W(t^{-1}(y))|\frac{dt^{-1}(y)}{dy}|.$$

[4]. If $W_1, W_2, \ldots, W_n$ are iid with pdf $f_W(w)$, assume that the MLE of $\hat{\boldsymbol{\theta}}$ is $\boldsymbol{\theta}_W(\boldsymbol{w})$ where the w_i are the observed values of W_i and $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)^T$.

If $Y_1, Y_2, \ldots, Y_n$ are iid and the y_i are the observed values of Y_i, then the likelihood is

$$L_Y(\boldsymbol{\theta}) = \left(\prod_{i=1}^{n} |\frac{dt^{-1}(y_i)}{dy}|\right) \prod_{i=1}^{n} f_W(t^{-1}(y_i)|\boldsymbol{\theta}) = c \prod_{i=1}^{n} f_W(t^{-1}(y_i)|\boldsymbol{\theta})$$

Hence the log likelihood is

$$\log(L_Y(\boldsymbol{\theta})) = d + \sum_{i=1}^{n} \log[f_W(t^{-1}(y_i)|\boldsymbol{\theta})] = d + \sum_{i=1}^{n} \log[f_W(w_i|\boldsymbol{\theta})] = d + \log[L_W(\boldsymbol{\theta})]$$

where $w_i = t^{-1}(y_i)$. Hence maximizing the $\log(L_Y(\boldsymbol{\theta}))$ is equivalent to maximizing $\log(L_W(\boldsymbol{\theta}))$ and

$$\hat{\boldsymbol{\theta}}_Y(\boldsymbol{y}) = \hat{\boldsymbol{\theta}}_W(\boldsymbol{w}) = \hat{\boldsymbol{\theta}}_W(t^{-1}(y_1), t^{-1}(y_2), \ldots, t^{-1}(y_n)). \tag{1.2}$$

Compare Meeker and Escobar (1998, p. 175).) [11]

This result is useful since if the MLE based on the W_i has simple inference, then the MLE based on the Y_i will also have simple inference. For example, If $W_1, W_2, \ldots, W_n$ are iid $\sim EXP(\theta = \log(\sigma), \lambda)$ and $Y_1, Y_2, \ldots, Y_n$ are iid Pareto $(\sigma = e^\theta, \lambda)$, then $Y = e^W = t(W)$ and $W = \log(Y) = t^{-1}(Y)$. The MLE of (θ, λ) based on the W_i is $(\hat{\theta}, \hat{\lambda}) = (W_{(1)}, \overline{W} - W_{(1)})$. Hence by (1.2) and invariance, the MLE of (σ, λ) based on the Y_i is $\hat{\sigma} = \exp(\hat{\theta}) = \exp(W_{(1)}) = Y_{(1)}$ and

$$\hat{\lambda} = \overline{W} - W_{(1)} = \frac{1}{n}\sum_{i=1}^{n} \log(Y_i) - \log(Y_{(1)}).$$

1.2 DISSERTATION OVERVIEW

The Dissertation is organized as follows. The first introductory chapter introduces the new transformed distributions and gives a review of the literature. Chapter 2 studies inference in both the exponential distribution and the half-normal distribution, we give a modified confidence interval for μ in the half-normal distribution which is better than the confidence interval given by Pewsey [15]. The Burr, largest extreme value, Pareto, power, Rayleigh, smallest extreme value, truncated extreme value, and Weibull distributions are obtained by transforming the exponential distribution. By applying the same transformation to the half normal distribution, new competitors for these distributions are obtained. We studied each of these sixteen transformed distributions in this chapter by allocating one section for each distribution. In each section we tried to give the pdf of the distribution, and its graph for selected values of it's parameter(s). The MLEs and confidence intervals for the parameter(s) were given for several distributions.

In chapter three, we present the results obtained from simulation studies to establish the actual coverage of the confidence intervals presented in chapter two. Sample sizes used in the simulations ranges from 5 to 500 and the number of runs ranges from 100 to 5000. The results of the simulation studies are found to give support for the confidence intervals presented in chapter two.

1.3 LITERATURE REVIEW

If Y has a (two parameter) exponential distribution, $Y \sim EXP(\theta, \lambda)$ then the probability density function (pdf) of Y is

$$f(y) = \frac{1}{\lambda} \exp\left(- \frac{(y-\theta)}{\lambda} \right) I[y \geq \theta]$$

where $\lambda > 0$ and θ is real. The cdf of Y is

$$F(y) = 1 - \exp\left(- \frac{(y-\theta)}{\lambda} \right), y \geq \theta$$

This is a location–scale family. If $X \sim EXP(\lambda)$, then $X \sim EXP(0, \lambda)$ has a one parameter exponential distribution and $X \sim G(1, \lambda)$ where G stands for the Gamma distribution. Inference for this distribution is discussed in Johnson and Kotz (1970, p. 219)[7], Mann, Schafer, and Singpurwalla (1974, p. 176)[10], Bury[3], Evans[6], Lehmann[9], and Krishnamorthy[8].

If Y has a half normal distribution, $Y \sim HN(\mu, \sigma^2)$, then the pdf of Y is

$$f(y) = \frac{2}{\sqrt{2\pi}\ \sigma} \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right)$$

where $\sigma > 0$ and $y \geq \mu$ and μ is real. This is a location-scale family. Let $\Phi(y)$ denote the standard normal cdf. Then the cdf of Y is

$$F(y) = 2\Phi\left(\frac{y-\mu}{\sigma}\right) - 1$$

for $y > \mu$ and $F(y) = 0$, otherwise. Inference for the this distribution is discussed by Pewsey [15].

CHAPTER 2

STATISTICAL DISTRIBUTIONS

2.1 THE (TWO PARAMETER) EXPONENTIAL DISTRIBUTION

If Y has a (two parameter) exponential distribution, $Y \sim EXP(\theta, \lambda)$ then the probability density function (pdf) of Y is

$$f(y) = \frac{1}{\lambda} \exp\left(-\frac{(y-\theta)}{\lambda}\right) I[y \geq \theta]$$

where $\lambda > 0$ and θ is real. The cdf of Y is

$$F(y) = 1 - \exp\left(-\frac{(y-\theta)}{\lambda}\right), y \geq \theta$$

This is a location–scale family. If $X \sim EXP(\lambda)$, then $X \sim EXP(0, \lambda)$ has a one parameter exponential distribution and $X \sim G(1, \lambda)$ where G stands for the Gamma distribution. Inference for this distribution is discussed in Johnson and Kotz (1970, p. 219) and Mann, Schafer, and Singpurwalla (1974, p. 176).

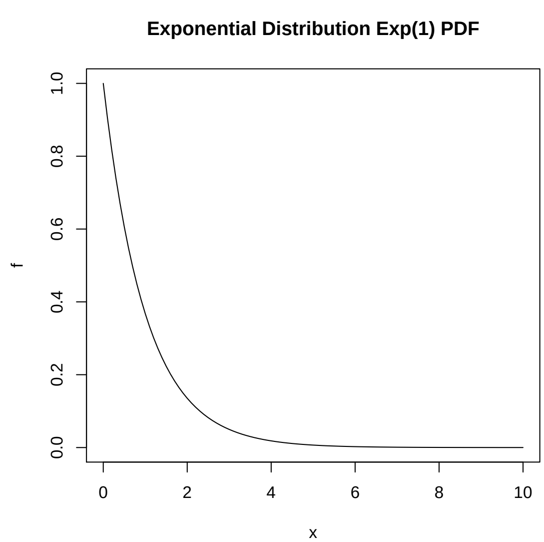

Figure 2.1. Plot of the pdf of the Exponential Distribution

Let $Y_1, ..., Y_n$ be iid $EXP(\theta, \lambda)$ random variables. Let $Y_{(1)} = \min(Y_1, ..., Y_n)$. Then the MLE

$$(\hat{\theta}, \hat{\lambda}) = \left(Y_{(1)}, \frac{1}{n} \sum_{i=1}^{n} (Y_i - Y_{(1)}) \right) = (Y_{(1)}, \overline{Y} - Y_{(1)}).$$

Let $D_n = n\hat{\lambda}$. For $n > 1$, an exact $100(1-\alpha)\%$ confidence interval (CI) for θ is

$$(Y_{(1)} - \hat{\lambda}[(\alpha)^{-1/(n-1)} - 1], Y_{(1)}) \tag{2.1}$$

while a $100(1-\alpha)\%$ CI for λ is

$$\left(\frac{2D_n}{\chi^2_{2(n-1),1-\alpha/2}}, \frac{2D_n}{\chi^2_{2(n-1),\alpha/2}} \right). \tag{2.2}$$

where $P(X < \chi^2_{n,\alpha}) = \alpha$ if X is chi-square with n degrees of freedom.

Let $T_n = \sum_{i=1}^n (Y_i - \theta) = n(\overline{Y} - \theta)$. If θ is known, then

$$\hat{\lambda}_\theta = \frac{\sum_{i=1}^n (Y_i - \theta)}{n} = \overline{Y} - \theta$$

is the uniformly minimum variance unbiased estimator (UMVUE) and maximum likelihood estimator (MLE) of λ, and a $100(1-\alpha)\%$ CI for λ is

$$\left(\frac{2T_n}{\chi^2_{2n,1-\alpha/2}}, \frac{2T_n}{\chi^2_{2n,\alpha/2}} \right). \tag{2.3}$$

Proof:

Let $X_i = Y_i - \theta \Rightarrow X_1, ..., X_n$ are iid EXP(λ) random variables. Then

$X_1, ..., X_n$ are iid $Gamma(1, \lambda)$

$\Rightarrow U = \sum_{i=1}^n X_i = \sum_{i=1}^n (Y_i - \theta) \sim Gamma(n, \lambda)$

$\Rightarrow W = 2U = 2\sum_{i=1}^n (Y_i - \theta) \sim Gamma(n, 2\lambda)$

$\Rightarrow V = W/\lambda \sim Gamma(n, 2) \sim \chi^2_{2n}$

$\Rightarrow 1 - \alpha = P(\chi^2_{2n,\alpha/2} < V < \chi^2_{2n,1-\alpha/2})$

$= P(\chi^2_{2n,\alpha/2} < \frac{2\sum_{i=1}^n (Y_i - \theta)}{\lambda} < \chi^2_{2n,1-\alpha/2})$

$= P(\frac{1}{\chi^2_{2n,1-\alpha/2}} < \frac{\lambda}{2\sum_{i=1}^n (Y_i - \theta)} < \frac{1}{\chi^2_{2n,\alpha/2}})$

$= P(\frac{2\sum_{i=1}^n (Y_i-\theta)}{\chi^2_{2n,1-\alpha/2}} < \lambda < \frac{2\sum_{i=1}^n (Y_i-\theta)}{\chi^2_{2n,\alpha/2}})$.

Using $\chi^2_{n,\alpha}/\sqrt{n} \approx \sqrt{2}z_\alpha + \sqrt{n}$, it can be shown that $\sqrt{n}$ CI length converges to $\lambda(z_{1-\alpha/2} - z_{\alpha/2})$ for CIs (2.2) and (2.3) (in probability). It can be shown that n length CI (2.1) converges to $-\lambda \log(\alpha)$ [12]. Proof:

$\chi^2_{2n,\alpha/2}/\sqrt{2n} \approx \sqrt{2}z_{\alpha/2} + \sqrt{2n}$..............$(A)$

$\Rightarrow \chi^2_{2n,\alpha/2}/\sqrt{n} = \sqrt{2}\chi^2_{2n,\alpha/2}/\sqrt{2n} \approx \sqrt{2}[\sqrt{2}z_{\alpha/2} + \sqrt{2n}]$

$= 2z_{\alpha/2} + 2\sqrt{n} = 2(z_{\alpha/2} + \sqrt{n})$..............$(B)$

Now

$$\sqrt{n}\text{ CI length } = \sqrt{n}(\frac{2\sum_{i=1}^n (Y_i-\theta)}{\chi^2_{2n,\alpha/2}} - \frac{2\sum_{i=1}^n (Y_i-\theta)}{\chi^2_{2n,1-\alpha/2}})$$

$$= (\frac{2\sum_{i=1}^n (Y_i-\theta)}{\chi^2_{2n,\alpha/2}/\sqrt{n}} - \frac{2\sum_{i=1}^n (Y_i-\theta)}{\chi^2_{2n,1-\alpha/2}/\sqrt{n}}) \text{ and by (B) above, this equals:}$$

$$(\frac{2\sum_{i=1}^n (Y_i-\theta)}{2(z_{\alpha/2}+\sqrt{n})} - \frac{2\sum_{i=1}^n (Y_i-\theta)}{2(z_{1-\alpha/2}+\sqrt{n})})$$

$$= \frac{\sum_{i=1}^n (Y_i-\theta)[z_{1-\alpha/2} - z_{\alpha/2}]}{(z_{\alpha/2}+\sqrt{n})(z_{1-\alpha/2}+\sqrt{n})} \frac{1/n}{1/n}$$

$$= \frac{1}{n}\frac{\sum_{i=1}^n (Y_i-\theta)[z_{1-\alpha/2} - z_{\alpha/2}]}{a_n} \xrightarrow{D} \lambda[z_{1-\alpha/2} - z_{\alpha/2}].$$

since

$$a_n \xrightarrow{D} 1.$$

Also

$$\lim n[\alpha^{-1/n-1} - 1] = \lim \frac{\alpha^{-1/n-1} - 1}{1/n} = \frac{0}{0}$$

so by L'Hopital Rule the limit equals

$$\lim \frac{\alpha^{-1/n-1}\log(\alpha)(-1/n^2)}{-1/n^2}$$

$$= -\log(\alpha) \lim \alpha^{-1/n} = -\log(\alpha).$$

Hence, n length CI (2.1) converges to

$$\lim n\hat{\lambda}[\alpha^{-1/n-1} - 1] = -\lambda \log(\alpha).$$

2.2 THE HALF NORMAL DISTRIBUTION

If Y has a half normal distribution, $Y \sim HN(\mu, \sigma^2)$, then the pdf of Y is

$$f(y) = \frac{2}{\sqrt{2\pi}\ \sigma} \exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right)$$

where $\sigma > 0$ and $y \geq \mu$ and μ is real. This is a location-scale family. Let $\Phi(y)$ denote the standard normal cdf. Then the cdf of Y is

$$F(y) = 2\Phi\left(\frac{y-\mu}{\sigma}\right) - 1$$

for $y > \mu$ and $F(y) = 0$, otherwise.

$E(Y) = \mu + \sigma\sqrt{2/\pi} \approx \mu + 0.797885\sigma.$

$VAR(Y) = \frac{\sigma^2(\pi-2)}{\pi} \approx 0.363380\sigma^2.$

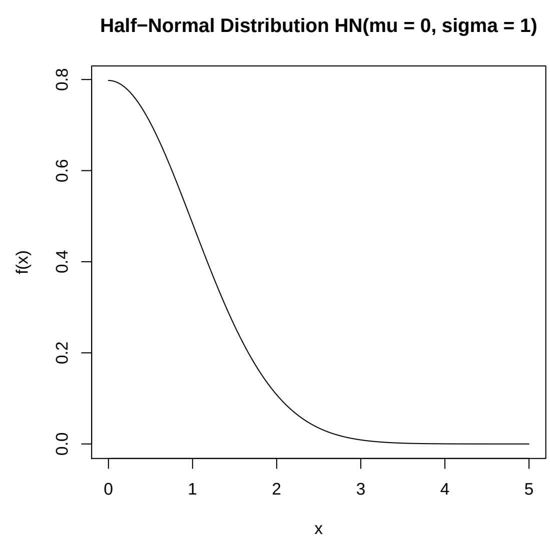

Figure 2.2. Plot of the pdf of the Half-Normal Distribution

This is an asymmetric location–scale family that has the same distribution as $\mu + \sigma|Z|$ where $Z \sim N(0,1)$. Note that $Z^2 \sim \chi^2_1$. Hence the formula for the rth moment of the χ^2_1 random variable can be used to find the moments of Y [12].

$MED(Y) = \mu + 0.6745\sigma$.

$MAD(Y) = 0.3990916\sigma$.

Notice that

$$f(y) = \frac{2}{\sqrt{2\pi}\ \sigma} I(y > \mu) \exp\left[(\frac{-1}{2\sigma^2})(y-\mu)^2 \right]$$

is a **1P–REF** if μ is known. Hence $\Theta = (0, \infty)$, $\eta = -1/(2\sigma^2)$ and $\Omega = (-\infty, 0)$.

$W = (Y-\mu)^2 \sim G(1/2, 2\sigma^2)$. If $Y_1, ..., Y_n$ are iid $HN(\mu, \sigma^2)$, then

$$T_n = \sum (Y_i - \mu)^2 \sim G(n/2, 2\sigma^2).$$

If μ is known, then the likelihood

$$L(\sigma^2) = c\ \frac{1}{\sigma^n} \exp\left[(\frac{-1}{2\sigma^2}) \sum (y_i - \mu)^2 \right],$$

and the log likelihood

$$\log(L(\sigma^2)) = d - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum (y_i-\mu)^2.$$

Hence

$$\frac{d}{d(\sigma^2)} \log(L(\sigma^2)) = \frac{-n}{2(\sigma^2)} + \frac{1}{2(\sigma^2)^2}\sum (y_i-\mu)^2 \overset{set}{=} 0,$$

or $\sum (y_i - \mu)^2 = n\sigma^2$ or

$$\hat{\sigma}^2 = \frac{1}{n}\sum (Y_i - \mu)^2.$$

Notice that

$$\frac{d^2}{d(\sigma^2)^2} \log(L(\sigma^2)) =$$

$$\frac{n}{2(\sigma^2)^2} - \left.\frac{\sum (y_i-\mu)^2}{(\sigma^2)^3}\right|_{\sigma^2 = \hat{\sigma}^2} = \frac{n}{2(\hat{\sigma}^2)^2} - \frac{n\hat{\sigma}^2}{(\hat{\sigma^2})^3}\frac{2}{2} = \frac{-n}{2\hat{\sigma}^2} < 0.$$

Thus $\hat{\sigma}^2$ is the UMVUE and MLE of σ^2 if μ is known, while $\hat{\mu} = Y_{(1)}$ is the ML estimate of μ for the case where σ is known.

Likelihood Based Confidence Intervals

Let $Y \sim HN(\mu, \sigma^2)$,$Y = \mu + \sigma X, X = |Z|, Z \sim \mathrm{N}(0, 1)$

Since $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (Y_i - \mu)^2}{n}$, then $n\hat{\sigma}^2 = \sum_{i=1}^n (Y_i - \mu)^2 = \sum_{i=1}^n (\sigma X_i)^2 \Rightarrow \frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_n$.

$\hat{\sigma}^2$ is a consistent estimator of σ^2 if μ is known and the $100(1-\alpha)\%$ confidence interval for σ^2 is

$$\left(\frac{T_n}{\chi^2_{n,1-\frac{\alpha}{2}}}, \frac{T_n}{\chi^2_{n,\frac{\alpha}{2}}}\right) \tag{2.4}$$

where $T_n = \sum_{i=1}^n (Y_i - \mu)^2$. If μ is unknown, let $D_n = \sum_{i=1}^n (Y_i - Y_{(1)})^2$. Then a $100(1-\alpha)\%$ large sample confidence interval for σ^2 is

$$\left(\frac{D_n}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{D_n}{\chi^2_{n-1,\frac{\alpha}{2}}}\right) \tag{2.5}$$

see corollary 2.1 below.

Pewsey [15] states that the limiting distribution of the MLE $\hat{\mu} = Y_{(1)}$ is

$$(\hat{\mu} - \mu)/\left[\sigma\Phi^{-1}\left(\tfrac{1}{2} + \tfrac{1}{2n}\right)\right] \xrightarrow{D} EXP(1)$$

where $\Phi(\cdot)$ denotes the distribution function of the standard normal distribution. It follows, therefore, that $\hat{\mu}$ is a consistent estimator for μ. An approximation to $\Phi^{-1}(\frac{1}{2} + \frac{1}{2n})$, based on a first order Taylor series expansion of the standard normal density, is given by $(\pi/2)^{1/2}/n$. This approximation is accurate to 2 decimal places for $n = 10$, and to 5 decimal places for $n = 50$ (Ref. [15], p. 1048).

An approximate $100(1-\alpha)\%$ confidence interval for μ is given by

$$\left(\hat{\mu} + \hat{\sigma}\log\left(\frac{\alpha}{2}\right)\Phi^{-1}\left(\frac{1}{2} + \frac{1}{2n}\right), \hat{\mu} + \hat{\sigma}\log\left(1 - \frac{\alpha}{2}\right)\Phi^{-1}\left(\frac{1}{2} + \frac{1}{2n}\right)\right) \tag{2.6}$$

Proof:

Let $f(y) = \exp(-y), y \geq 0$, and $F(y) = 1 - \exp(-y)$, $y \geq 0$, then

$\frac{\alpha}{2} = F(y_{\frac{\alpha}{2}}) = 1 - \exp(-y_{\frac{\alpha}{2}})$

$\Rightarrow \exp(-y_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} \Rightarrow y_{\frac{\alpha}{2}} = -\log(1 - \frac{\alpha}{2})$, and

$1 - \frac{\alpha}{2} = F(y_{1-\frac{\alpha}{2}}) = 1 - \exp(-y_{1-\frac{\alpha}{2}})$

$\Rightarrow \frac{\alpha}{2} = \exp(-y_{1-\frac{\alpha}{2}}) \Rightarrow -y_{1-\frac{\alpha}{2}} = \log(\frac{\alpha}{2})$ or $y_{1-\frac{\alpha}{2}} = -\log(\frac{\alpha}{2})$.

$\Rightarrow 1 - \alpha \approx P(y_{\frac{\alpha}{2}} < \frac{Y_{(1)} - \mu}{\sigma\Phi^{-1}(\frac{1}{2} + \frac{1}{2n})} < y_{1-\frac{\alpha}{2}})$

$\Rightarrow 1 - \alpha \approx P(-\log(1 - \frac{\alpha}{2}) < \frac{Y_{(1)} - \mu}{\sigma\Phi^{-1}(\frac{1}{2} + \frac{1}{2n})} < -\log(\frac{\alpha}{2})) \approx$

$P(\log(\frac{\alpha}{2}) < \frac{\mu - Y_{(1)}}{\sigma\Phi^{-1}(\frac{1}{2} + \frac{1}{2n})} < \log(1 - \frac{\alpha}{2}))$

$\Rightarrow P(Y_{(1)} + \sigma\Phi^{-1}(\frac{1}{2} + \frac{1}{2n})\log(\frac{\alpha}{2}) < \mu < Y_{(1)} + \sigma\Phi^{-1}(\frac{1}{2} + \frac{1}{2n})\log(1 - \frac{\alpha}{2})) \approx 1 - \alpha$.

Examining (2.1), we suggest that a new and better CI is

$$(\hat{\mu} + \hat{\sigma}\log(\alpha)\Phi^{-1}(\frac{1}{2} + \frac{1}{2n})(1 + \frac{13}{n^2}), \hat{\mu}). \tag{2.7}$$

If σ is known then a large-sample confidence interval for μ with the same nominal confidence level is obtained by substituting σ for $\hat{\sigma}$ in (2.7). Pewsey (2002, p. 1048) said that

$$D_n \xrightarrow{D} \chi^2_{n-1}.$$

This can't happen as since the righthand limit depends on n, hence we introduce the following theorem.

Theorem 2.1. Let $T_n = \sum_{i=1}^n (Y_i - \mu)^2$ and $D_n = \sum_{i=1}^n (Y_i - Y_{(1)})^2$.

Then

$$D_n - T_n \xrightarrow{D} -\sigma^2\chi^2_2.$$

Proof.

$D_n = \sum_{i=1}^n (Y_i - \mu + \mu - Y_{(1)})^2$

$= \sum_{i=1}^n (Y_i - \mu)^2 + 2\sum_{i=1}^n (Y_i - \mu)(\mu - Y_{(1)}) + \sum_{i=1}^n (\mu - Y_{(1)})^2$

$= T_n + n(\mu - Y_{(1)})^2 - 2(Y_{(1)} - \mu)\sum_{i=1}^{n}(Y_i - \mu).$

then $\frac{D_n}{\sigma^2} = \frac{T_n}{\sigma^2} + \frac{1}{n}\frac{1}{\sigma^2}[n(Y_{(1)} - \mu)]^2 - 2[\frac{n(Y_{(1)}-\mu)}{\sigma}][\frac{\sum_{i=1}^{n}(Y_i-\mu)}{n\sigma}]$

$\Rightarrow$

$$\frac{D_n - T_n}{\sigma^2} = \frac{1}{n}[\frac{n(Y_{(1)} - \mu)}{\sigma}]^2 - 2[\frac{n(Y_{(1)} - \mu)}{\sigma}][\frac{\sum_{i=1}^{n}(Y_i - \mu)}{n\sigma}] \tag{2.8}$$

Pewsey (2002, p. 1048) showed that

$$\frac{Y_{(1)}-\mu}{\left[\sigma\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right)\right]} \xrightarrow{D} EXP(1).$$

and since

$$\frac{\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right)}{\sqrt{\frac{\pi}{2}}/n} \xrightarrow{D} 1,$$

then, by Slutsky's Theorem

$$\frac{Y_{(1)}-\mu}{\left[\sigma\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right)\right]}\frac{\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right)}{\sqrt{\frac{\pi}{2}}/n} \xrightarrow{D} EXP(1).$$

Hence

$$\frac{n(Y_{(1)}-\mu)}{\sqrt{\frac{\pi}{2}}\sigma} \xrightarrow{D} EXP(1).$$

By the law of large numbers, the third term

$$\overline{Z} = \frac{\sum_{i=1}^{n}(Y_i-\mu)}{n\sigma} \xrightarrow{D} \text{E(Z)} = \sqrt{\frac{2}{\pi}}$$

where $Z_i = \frac{Y_i-\mu}{\sigma} \sim HN(0,1)$.

Since

$$\frac{n(Y_{(1)}-\mu)}{\sigma} \xrightarrow{D} \sqrt{\frac{\pi}{2}}EXP(1),$$

$$\frac{1}{n}[\frac{n(Y_{(1)}-\mu)}{\sigma}]^2 \xrightarrow{D} 0.$$

Hence

$$\frac{D_n - T_n}{\sigma^2} \xrightarrow{D} 0 - 2\sqrt{\frac{\pi}{2}}EXP(1)\sqrt{\frac{2}{\pi}} = -2\text{EXP}(1).$$

Or

$$D_n - T_n \xrightarrow{D} -\sigma^2\chi_2^2. \; QED$$

Let $V_n = \sigma^2\chi_2^2$ and $T_{n-p} = \sum_{i=1}^{n-p}(Y_i - \mu)^2$. Then

$$D_n = T_{n-p} + \sum_{i=n-p+1}^{n}(Y_i - \mu)^2 - V_n$$

where

$$\frac{V_n}{\sigma^2} \xrightarrow{D} \chi_2^2.$$

Hence

$$\frac{D_n}{T_{n-p}} = 1 + \frac{\sum_{i=n-p+1}^{n}(Y_i - \mu)^2}{T_{n-p}} - \frac{V_n}{T_{n-p}}$$

Hence

$$\frac{D_n}{T_{n-p}} \xrightarrow{D} 1.$$

Since

$$\frac{T_{n-p}}{\sigma^2} = \sum_{i=1}^{n-p}(\frac{Y_i - \mu)}{\sigma})^2 \sim \chi_{n-p}^2,$$

$$\frac{D_n}{\sigma^2} \approx \chi_{n-p}^2$$

where p is a nonnegative integer. Pewsey (2002) used $p = 1$.

Corollary 2.1. If μ is known, then a $100(1-\alpha)\%$ confidence interval for σ is:

$$(\sqrt{\frac{T_n}{\chi^2_{n,1-\frac{\alpha}{2}}}}, \sqrt{\frac{T_n}{\chi^2_{n,\frac{\alpha}{2}}}}) \tag{2.9}$$

and if μ is unknown, then a large sample $100(1-\alpha)\%$ confidence interval for σ is

$$\left(\sqrt{\frac{D_n}{\chi^2_{n-1,1-\frac{\alpha}{2}}}}, \sqrt{\frac{D_n}{\chi^2_{n-1,\frac{\alpha}{2}}}}\right). \tag{2.10}$$

Using $\chi^2_{n,\alpha}/\sqrt{n} \approx \sqrt{2}Z_\alpha + \sqrt{n}$, it can be shown that $\sqrt{n}CI$ length converges in probability to $\sqrt{2}\sigma^2[Z_{1-\alpha/2} - Z_{\alpha/2}]$ for CIs (2.4) and (2.5). Also it can be shown that nCI length converges to $-\sigma\log(\alpha)\sqrt{\pi/2}$ for CI (2.7).

Proof:

$$\begin{aligned}
\sqrt{n}CI \text{ length } &= \sqrt{n}\left(\frac{T_n}{\chi^2_{n,\alpha/2}} - \frac{T_n}{\chi^2_{n,1-\alpha/2}}\right) \\
&= \left(\frac{\sum_{i=1}^n (Y_i-\mu)^2}{\chi^2_{n,\alpha/2}/\sqrt{n}} - \frac{\sum_{i=1}^n (Y_i-\mu)^2}{\chi^2_{n,1-\alpha/2}/\sqrt{n}}\right) \\
&= \left(\frac{\sum_{i=1}^n (Y_i-\mu)^2}{\sqrt{2}z_{\alpha/2}+\sqrt{n}} - \frac{\sum_{i=1}^n (Y_i-\mu)^2}{\sqrt{2}z_{1-\alpha/2}+\sqrt{n}}\right)\frac{1/n}{1/n}
\end{aligned}$$

$$= \frac{1}{n}\frac{\sqrt{2}\sum_{i=1}^n (Y_i-\mu)^2[z_{1-\alpha/2}-z_{\alpha/2}]}{a_n} \quad \xrightarrow{D} \quad \sqrt{2}\sigma^2[z_{1-\alpha/2} - z_{\alpha/2}]$$

since

$$a_n \quad \xrightarrow{D} \quad 1.$$

Now nCI length of (2.7)

$$= -n[\hat{\sigma}\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right)\log\alpha]$$

$$\approx -\sigma\log\alpha\sqrt{\pi/2}$$

since a Taylor series approximation of $\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right)$ is $(\pi/2)^{1/2}/n$.

2.3 THE BURR DISTRIBUTION

If $Y \sim Burr(\phi, \lambda)$, then the cdf of Y is $F(y) = 1 - \exp(\frac{-\log(1+y^{\phi})}{\lambda}) = 1 - (1 + y^{\phi})^{(-\frac{1}{\lambda})}$ for $y > 0$ and the pdf of Y is

$f(y) = \frac{1}{\lambda}\frac{\phi y^{\phi-1}}{(1+y^{\phi})^{1+\frac{1}{\lambda}}}, y > 0, \phi > 0, \lambda > 0.$

Let $W = \log(1 + Y^{\phi})$, then

$P(W \leq w) = P(\log(1 + Y^{\phi}) \leq w) = P(1 + Y^{\phi} \leq e^{w}) = P(Y \leq (e^{w} - 1)^{\frac{1}{\phi}})$

$= 1 - (1 + ((e^{w} - 1)^{\frac{1}{\phi}})^{\phi})^{-\frac{1}{\lambda}} = 1 - [1 + e^{w} - 1]^{-\frac{1}{\lambda}} = 1 - e^{-\frac{w}{\lambda}}.$

Hence $W \sim$ EXP(λ).

Let $Y = (e^{W} - 1)^{\frac{1}{\phi}}$. Then

$F_Y(y) = P(Y \leq y) = P((e^{W} - 1)^{\frac{1}{\phi}} \leq y) = P(e^{W} - 1 \leq y^{\phi}) = P(e^{W} \leq 1 + y^{\phi}) = P(W \leq \log(1 + y^{\phi})) = 1 - \exp(\frac{-\log(1+y^{\phi})}{\lambda}).$

Hence $Y \sim Burr(\phi, \lambda)$.

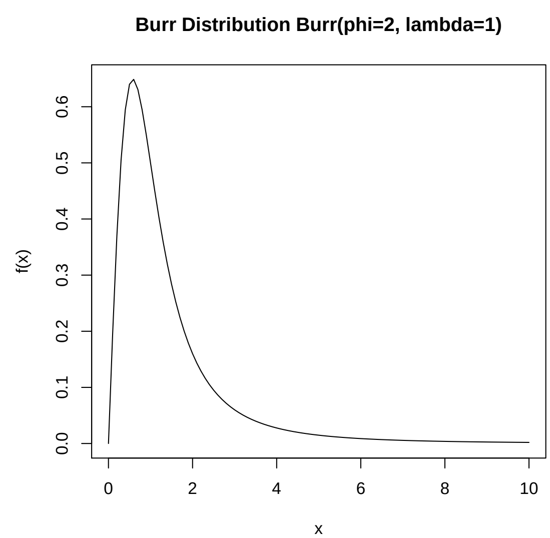

Figure 2.3. Plot of the pdf of the Burr Distribution

2.4 THE HBURR DISTRIBUTION

If $Y \sim HBurr(\phi, \lambda)$, then

$f(y) = \frac{2}{\sqrt{2\pi}\lambda} \frac{\phi y^{\phi-1}}{y^{\phi}+1} \exp(\frac{-(\log(y^{\phi}+1))^2}{2\lambda^2}), y > 0, , \phi > 0, \lambda > 0.$

If W has a half normal distribution, $W \sim HN(0, \lambda)$,

let $Y = (e^W - 1)^{\frac{1}{\phi}}$. Then $Y^{\phi} = e^W - 1 \Rightarrow e^W = Y^{\phi} + 1$

$\Rightarrow W = s(Y) = \log(Y^{\phi} + 1)$.

Then $f_Y(y) = g_W(s(y))|\frac{ds(y)}{dy}| = \frac{2}{\sqrt{2\pi}\lambda} \exp(\frac{-(\log(y^{\phi}+1)-0)^2}{2\lambda^2}) \frac{1}{y^{\phi}+1} \phi y^{\phi-1}$

$= \frac{2}{\sqrt{2\pi}\lambda} \exp(\frac{-(\log(y^{\phi}+1))^2}{2\lambda^2}) \frac{\phi y^{\phi-1}}{y^{\phi}+1}, y > 0,$

$\Rightarrow Y \sim HBurr(\phi, \lambda)$.

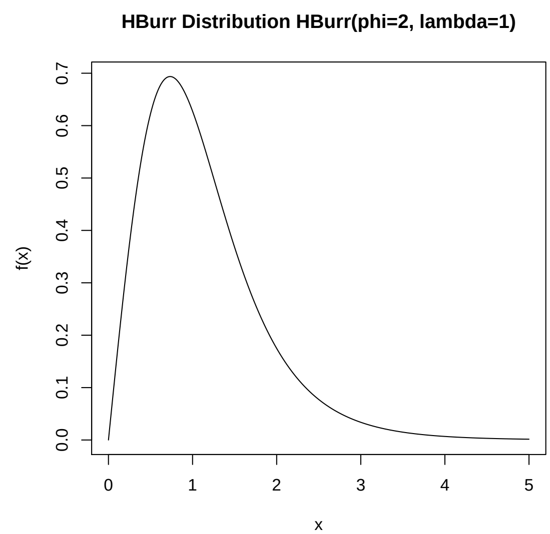

Figure 2.4. Plot of the pdf of the HBurr Distribution

Let $W = \log(1 + Y^{\phi})$, then $Y = r(W) = (e^W - 1)^{\frac{1}{\phi}}$

$\Rightarrow g_W(w) = f_Y(r(w))|\frac{dr(w)}{dw}| = \frac{2}{\sqrt{2\pi}\lambda} \exp(-\frac{w^2}{2\lambda^2}) \frac{\phi[(e^w-1)^{\frac{1}{\phi}}]^{\phi-1}}{[(e^w-1)^{\frac{1}{\phi}}]^{\phi}+1} |\frac{1}{\phi}(e^w - 1)^{\frac{1}{\phi}-1} e^w|$

$= \frac{2}{\sqrt{2\pi}\lambda} \exp(-\frac{w^2}{2\lambda^2}) \frac{\phi(e^w-1)(e^w-1)^{\frac{-1}{\phi}}}{(e^w-1)+1} |\frac{1}{\phi} \frac{(e^w-1)^{\frac{1}{\phi}} e^w}{e^w-1}|$

$= \frac{2}{\sqrt{2\pi}\lambda} \exp(-\frac{w^2}{2\lambda^2}) \frac{\phi(e^w - 1)}{(e^w-1)^{\frac{1}{\phi}} e^w} |\frac{1}{\phi}(e^w - 1)^{\frac{1}{\phi}-1} e^w|$

$= \frac{2}{\sqrt{2\pi}\lambda} e^{-\frac{w^2}{2\lambda^2}}$, $w \geq 0$.

$\Rightarrow W \sim HN(0, \lambda)$.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $HBurr(\phi, \lambda)$,and if ϕ is known, then the likelihood

$$L(\lambda) = \prod_{i=1}^{n} f(y_i) = (\frac{2}{\sqrt{2\pi}\lambda})^n \frac{\phi^n \prod_{i=1}^n y_i^{\phi-1}}{\prod_{i=1}^n (y_i^{\phi} + 1)} \exp(\sum_{i=1}^{n} \frac{-(\log(y_i^{\phi} + 1))^2}{2\lambda^2}),$$

and the log likelihood

$$\log L = n \log(2) - n \log(\sqrt{2\pi}) - n \log \lambda - \sum_{i=1}^{n} \frac{(\log(y_i^{\phi} + 1))^2}{2\lambda^2} + n \log \phi$$

$$+ \sum_{i=1}^{n} ((\phi - 1) \log y_i - \log(y_i^{\phi} + 1)).$$

Hence

$$\frac{d \log L}{d\lambda} = \frac{-n}{\lambda} + \frac{2 \sum_{i=1}^n (\log(y_i^{\phi} + 1))^2}{2\lambda^3} := 0$$

or

$$\lambda^2 = \frac{\sum_{i=1}^n (\log(y_i^{\phi} + 1))^2}{n}, or \hat{\lambda} = \sqrt{\frac{\sum_{i=1}^n (\log(y_i^{\phi} + 1))^2}{n}}.$$

Notice that $\frac{d^2 \log L}{d\lambda^2}|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \frac{3 \sum_{i=1}^n (\log(y_i^{\phi}+1))^2}{\hat{\lambda}^4} = \frac{n}{\hat{\lambda}^2}[1 - 3] < 0$. Hence $\hat{\lambda}$ is the MLE of λ if ϕ is known.

2.5 THE LARGEST EXTREME VALUE DISTRIBUTION

If Y has a Largest Extreme Value, $Y \sim LEV(\theta, \sigma)$,

then the pdf of Y is

$$f(y) = \frac{1}{\sigma} \exp(-(\frac{y-\theta}{\sigma})) \exp[-\exp(-(\frac{y-\theta}{\sigma}))]$$

where y and θ are real and $\sigma > 0$. This distribution is a location scale family. The cdf of Y is

$$F(y) = \exp[-\exp(-(\frac{y-\theta}{\sigma}))].$$

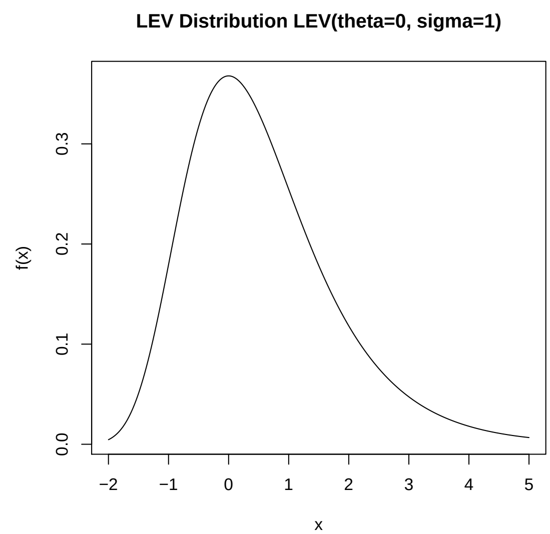

Figure 2.5. Plot of the pdf of the Largest Extreme Value Distribution

If $W \sim EXP(1)$, let $Y = -\sigma \log W + \theta$. Then

$F_Y(y) = P(Y \leq y) = P(-\sigma \log W + \theta \leq y) = P(\log W \geq -\frac{y-\theta}{\sigma})$

$= P(W \geq \exp(-\frac{y-\theta}{\sigma})) = 1 - [1 - \exp(-\exp(-\frac{(y-\theta)}{\sigma}))] = \exp(-\exp(-\frac{(y-\theta)}{\sigma})),$

$\sigma \geq 0, -\infty < y < \infty$

$\Rightarrow Y \sim LEV(\theta, \sigma).$

Let $W = \exp(-\frac{(Y-\theta)}{\sigma})$. Then $F_W(w) = P(W \leq w) = P(\exp(-\frac{(Y-\theta)}{\sigma}) \leq w) =$

$P(-\frac{(Y-\theta)}{\sigma} \leq \log w) = P(Y - \theta \geq -\sigma \log w) = P(Y \geq \theta - \sigma \log w) = 1 - \exp(-\exp(-\frac{(\theta-\sigma \log w)-\theta}{\sigma})) = 1 - \exp(-w)$

$\Rightarrow W \sim EXP(1)$.

2.6 THE HLEV DISTRIBUTION

If $Y \sim HLEV(\theta, \lambda)$, then the pdf of Y is

$$f(y) = \frac{2}{\sqrt{2\pi}} \frac{1}{\lambda} \exp(-\frac{(y-\theta)}{\lambda}) \exp(-\frac{1}{2}[\exp(\frac{-(y-\theta)}{\lambda})]^2), y \in R,\ \theta \in R,\ \lambda > 0,$$

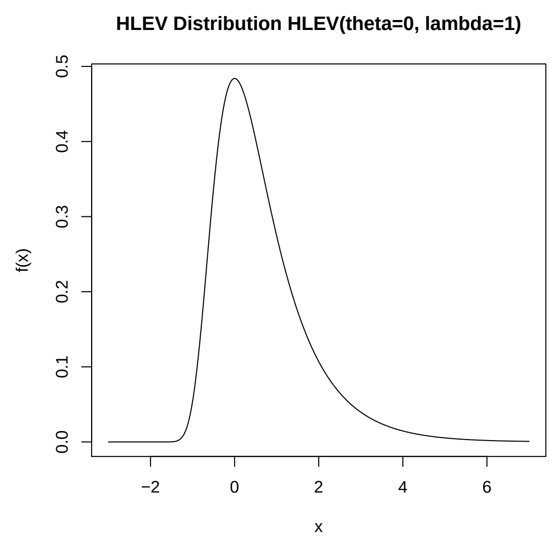

Figure 2.6. Plot of the pdf of the HLEV Distribution

If W has a half normal distribution, $W \sim HN(0,1)$,

then $g_W(w) = \frac{2}{\sqrt{2\pi}} \exp(\frac{-w^2}{2})$ for $w \geq 0$. Let $Y = -\lambda \log(W) + \theta$, then $W = s(Y) = \exp(-\frac{Y-\theta}{\lambda})$

$\Rightarrow f_Y(y) = g_W(s(y))|\frac{ds(y)}{dy}| = \frac{2}{\sqrt{2\pi}} \exp(\frac{-\exp(\frac{-2(y-\theta)}{\lambda})}{2})| \exp(-\frac{y-\theta}{\lambda})\frac{-1}{\lambda}|$

$= \frac{2}{\sqrt{2\pi}}\frac{1}{\lambda}\exp(-\frac{(y-\theta)}{\lambda})\exp(-\frac{1}{2}[\exp(\frac{-(y-\theta)}{\lambda})]^2), y \in R,\ \theta \in R,\ \lambda > 0,$

$\Rightarrow Y \sim HLEV(\theta, \lambda).$

Let $W = \exp(\frac{-(Y-\theta)}{\lambda})$, then $Y = r(W) = -\lambda\log(W) + \theta$

$\Rightarrow g_W(w) = f_Y(r(w))|\frac{dr(w)}{dw}|$

$= \frac{2}{\sqrt{2\pi}\lambda}\exp(-\frac{(-\lambda\log(w)+\theta-\theta)}{\lambda})\exp(-\frac{1}{2}[\exp(-\frac{(-\lambda\log(w)+\theta-\theta)}{\lambda}]^2).|\frac{-\lambda}{w}|$

$= \frac{2}{\sqrt{2\pi}\lambda}\exp(\log(w))\exp[-\frac{1}{2}w^2](\frac{\lambda}{w}) = \frac{2}{\sqrt{2\pi}\lambda}w\exp(-\frac{w^2}{2})\frac{\lambda}{w}$

$= \frac{2}{\sqrt{2\pi}}e^{-\frac{w^2}{2}},\ w \geq 0.$

$\Rightarrow W \sim HN(0, 1).$

2.7 THE PARETO DISTRIBUTION

If Y has a Pareto distribution, $Y \sim PAR(\sigma, \lambda)$, then the pdf of Y is

$$f(y) = \frac{\frac{1}{\lambda}\sigma^{1/\lambda}}{y^{1+1/\lambda}}$$

where $y \geq \sigma$, $\sigma > 0$, and $\lambda > 0$. The cdf of Y is $F(y) = 1 - (\sigma/y)^{1/\lambda}$ for $y > \sigma$.

Let $W = \log(Y)$, then

$P(W \leq w) = P(\log(Y) \leq w) = P(Y \leq e^w) = 1-(\frac{\sigma}{e^w})^{\frac{1}{\lambda}} = 1-(\sigma e^{-w})^{\frac{1}{\lambda}} = 1-\sigma^{\frac{1}{\lambda}}e^{\frac{-w}{\lambda}}.$

Hence

$f_W(w) = -\sigma^{\frac{1}{\lambda}}e^{\frac{-w}{\lambda}}\frac{-1}{\lambda} = \frac{1}{\lambda}\sigma^{\frac{1}{\lambda}}e^{\frac{-w}{\lambda}} = \frac{1}{\lambda}e^{\log\sigma^{\frac{1}{\lambda}}}e^{\frac{-w}{\lambda}} = \frac{1}{\lambda}e^{-\frac{(w-\log\sigma)}{\lambda}}, w \geq \log\sigma$

$\Rightarrow W \sim EXP(\theta = \log(\sigma), \lambda).$

If $W \sim EXP(\theta = \log(\sigma), \lambda)$, let $Y = e^W$. Then

$F_Y(y) = P(Y \leq y) = P(e^W \leq y) = P(W \leq \log y) = 1 - \exp(\frac{-(\log y - \log\sigma)}{\lambda})$

$= 1 - [e^{\frac{-\log y}{\lambda}}e^{\log(\sigma^{\frac{1}{\lambda}})}] = 1 - \sigma^{\frac{1}{\lambda}}e^{\frac{-\log y}{\lambda}} = 1 - \sigma^{\frac{1}{\lambda}}e^{\log(y^{\frac{-1}{\lambda}})} = 1 - \sigma^{\frac{1}{\lambda}}y^{\frac{-1}{\lambda}}$

$\Rightarrow f(y) = \frac{1}{\lambda}\sigma^{\frac{1}{\lambda}}y^{-\frac{1}{\lambda}-1} = \frac{\frac{1}{\lambda}\sigma^{\frac{1}{\lambda}}}{y^{1+\frac{1}{\lambda}}}, y \geq \sigma$

$\Rightarrow Y \sim PAR(\sigma, \lambda).$

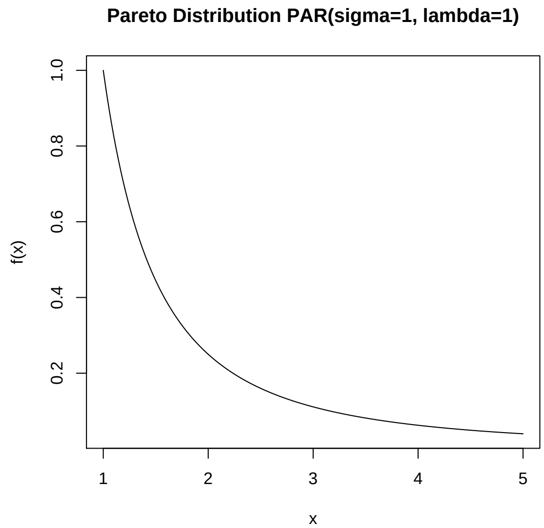

Figure 2.7. Plot of the pdf of the Pareto Distribution

Let $\theta = \log(\sigma)$. The MLE $(\hat{\theta}, \hat{\lambda}) = (W_{(1)}, \overline{W} - W_{(1)})$, and by invariance, the MLE $(\hat{\sigma}, \hat{\lambda}) = (e^{\hat{\theta}}, \overline{W} - W_{(1)}) = (e^{W_{(1)}}, \frac{1}{n}\sum_{i=1}^n \log Y_i - \log Y_{(1)}) = (Y_{(1)}, \frac{1}{n}\sum_{i=1}^n (\log Y_i - \log Y_{(1)})) = (Y_{(1)}, \frac{1}{n}\sum_{i=1}^n \log(Y_i/Y_{(1)}))$.

If σ is known,

$$\hat{\lambda} = \frac{\sum_{i=1}^n \log(Y_i/\sigma)}{n}$$

is the UMVUE and MLE of λ.

Inference is simple. If $\theta = \log(\sigma)$ so $\sigma = e^{\theta}$, then a 100 $(1-\alpha)$% CI for θ is (2.1). A 100 $(1-\alpha)$% CI for σ is obtained by exponentiating the endpoints of (2.1), and a 100 $(1-\alpha)$ % CI for λ is (2.2). Let $D_n = \sum_{i=1}^n (W_i - W_{(1)}) = n\hat{\lambda}$. For $n > 1$, a $100(1-\alpha)$% CI for θ is

$$(W_{(1)} - \hat{\lambda}[(\alpha)^{-1/(n-1)} - 1], W_{(1)}).$$

Exponentiate the endpoints for a $100(1-\alpha)$% CI for σ to get

$$(\exp(W_{(1)} - \hat{\lambda}[(\alpha)^{-1/(n-1)} - 1]), \exp(W_{(1)})). \tag{2.11}$$

A $100(1-\alpha)\%$ CI for λ is

$$\left(\frac{2D_n}{\chi^2_{2(n-1),1-\alpha/2}}, \frac{2D_n}{\chi^2_{2(n-1),\alpha/2}}\right). \tag{2.12}$$

These two exact CIs seem to be new.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $PAR(\sigma, \lambda)$, then the likelihood $L(\lambda) = \prod_{i=1}^n f(y_i) = \prod_{i=1}^n \frac{\frac{1}{\lambda}\sigma^{1/\lambda}}{y_i^{1+1/\lambda}} = (\frac{1}{\lambda})^n (\sigma^{\frac{1}{\lambda}})^n \prod_{i=1}^n \frac{1}{y_i^{1+1/\lambda}}$

and the log likelihood

$\log L = n\log(\lambda) + \frac{n}{\lambda}\log(\sigma) - \sum_{i=1}^n (1+\frac{1}{\lambda})\log y_i.$

Hence

$\frac{d\log L}{d\lambda} = -\frac{n}{\lambda} - \frac{n\log\sigma}{\lambda^2} - \sum_{i=1}^n \log y_i(-\frac{1}{\lambda^2}) = -\frac{n}{\lambda} - \frac{n\log\sigma}{\lambda^2} + \frac{1}{\lambda^2}\sum_{i=1}^n \log y_i := 0$

$\Rightarrow \frac{1}{\lambda}[\sum_{i=1}^n \log y_i - n\log\sigma] = n \Rightarrow \hat{\lambda} = \frac{\sum_{i=1}^n \log(\frac{y_i}{\hat{\sigma}})}{n}.$

Notice that

$$\frac{d^2\log L}{d\lambda^2}|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} + \frac{2n\log\sigma}{\hat{\lambda}^3} - \frac{2}{\hat{\lambda}^3}\sum_{i=1}^n \log y_i$$

$$= \frac{n}{\hat{\lambda}^2}(1) + \frac{2(\sum_{i=1}^n \log\sigma)n^3}{(\sum_{i=1}^n \log\frac{y_i}{\sigma})^3} - \frac{2n^3\sum_{i=1}^n \log y_i}{(\sum_{i=1}^n \log\frac{y_i}{\sigma})^3}$$

$$= \frac{n}{\lambda^2} + \frac{n^3}{(\sum_{i=1}^n \log\frac{y_i}{\sigma})^3}[2\sum_{i=1}^n \log\sigma - 2\sum_{i=1}^n \log y_i]$$

$$= \frac{n}{\lambda^2} + \frac{2n^3}{(\sum_{i=1}^n \log\frac{y_i}{\sigma})^3}(-\sum_{i=1}^n \log\frac{y_i}{\sigma}) = \frac{n}{\lambda^2} - 2\frac{n^3}{(\sum_{i=1}^n \log\frac{y_i}{\sigma})^2}$$

$$= \frac{n^2}{(\sum_{i=1}^n \log\frac{y_i}{\sigma})^2} - \frac{2n^3}{(\sum_{i=1}^n \log\frac{y_i}{\sigma})^2} < 0.$$

Hence $\hat{\lambda}$ is the MLE of λ given σ.

If neither σ nor λ are known, notice that

$$f(y) = \frac{1}{y}\frac{1}{\lambda}\exp[-(\frac{\log(y)-\log(\sigma)}{\lambda})]I(y_{(1)} \geq \sigma),$$

Hence the likelihood

$$L(\lambda, \sigma) = c\frac{1}{\lambda^n} \exp[-\sum_{i=1}^{n}(\frac{\log(y_i) - \log(\sigma)}{\lambda})]I(y_{(1)} \geq \sigma)$$

and the log likelihood is

$$\log L(\lambda, \sigma) = [d - n\log(\lambda) - \sum_{i=1}^{n}(\frac{\log(y_i) - \log(\sigma)}{\lambda})]I(y_{(1)} \geq \sigma)$$

Let $w_i = \log(y_i)$ and $\theta = \log(\sigma)$, so $\sigma = e^{\theta}$. Then the log likelihood is

$$\log L(\lambda, \sigma) = [d - n\log(\lambda) - \sum_{i=1}^{n}(\frac{w_i - \theta}{\lambda})]I(w_{(1)} \geq \theta),$$

which has the same form as the log likelihood of the EXP(θ, λ) distribution. Hence $(\hat{\lambda}, \hat{\theta}) = (\overline{W} - W_{(1)}, W_{(1)})$, and by invariance, the MLE

$$(\hat{\lambda}, \hat{\sigma}) = (\overline{W} - W_{(1)}, Y_{(1)}).$$

A second equation (corresponding to $d\log L/d\sigma = 0$) can not be obtained in the usual way since $\log L$ is unbounded on the random variable Y, $\log L$ must be maximized subject to the constraint:

$$\hat{\sigma} \leq minY_i.$$

By inspection, the value of $\hat{\sigma}$ which maximizes L is

$$\hat{\sigma} = minY_i = Y_{(1)}$$

so,

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} \log(\frac{Y_i}{Y_{(1)}})}{n}.$$

2.8 THE HPARETO DISTRIBUTION

If $Y \sim HPAR(\theta, \lambda)$, then

$f(y) = \frac{2}{\sqrt{2\pi}\lambda y} \exp(\frac{-(\log(y)-\log(\theta))^2}{2\lambda^2}), y \geq \theta,\ \lambda > 0,$ and $\theta > 0.$

This distribution is unimodal with the mode at $y = \theta$ and $f(\theta) = \frac{2}{\sqrt{2\pi}\lambda\theta}$

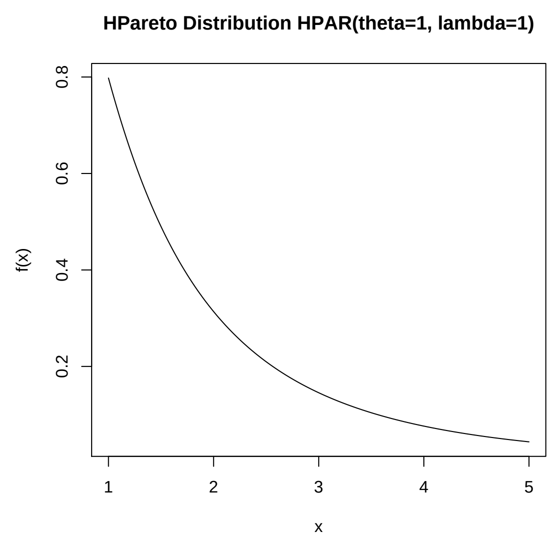

Figure 2.8. Plot of the pdf of the HPareto Distribution

Proof:

From the graph , the mode occurs at $Y_{(1)} = \theta$. Also from the formula of the pdf of Y we want to maximize $\frac{2}{\sqrt{2\pi}\lambda y} \exp(\frac{-(\log(y)-\log(\theta))^2}{2\lambda^2})$ and that happens at $Y_{(1)}$.

In addition, $\frac{df(y)}{dy}$ has no zeros on (θ, ∞). Proof:

$\log(f(y)) = c - \log y - \frac{(\log y - \log\theta)^2}{2\lambda^2}$

So $\frac{d}{dy}\log f(y) = 0 - \frac{1}{y} - \frac{1}{\lambda^2}(\log y - \log\theta)\frac{1}{y} = \frac{1}{y}[-1 - \frac{1}{\lambda^2}(\log(\frac{y}{\theta}))] := 0$

or

$1 = \frac{-1}{\lambda^2}\log(\frac{y}{\theta})$, or $y = \theta e^{-\lambda^2}$ which is not in support of Y.

If W has a half normal distribution, $W \sim HN(\mu, \sigma)$,

then $g_W(w) = \frac{2}{\sqrt{2\pi}\sigma} \exp(\frac{-(w-\mu)^2}{2\sigma^2}), w \geq \mu.$

Let $Y = e^W$, then $W = s(Y) = \log(Y)$

$\Rightarrow$

$$f_Y(y) = g_W(s(y))|\frac{ds(y)}{dy}| = \frac{2}{\sqrt{2\pi}\sigma}\exp(\frac{-(\log(y)-\mu)^2}{2\sigma^2})|\frac{1}{y}|$$

$(w \geq \mu \Rightarrow \log(y) \geq \mu \Rightarrow y \geq e^\mu \geq 0 \Rightarrow |y| = y)$

$$= \frac{2}{y\sqrt{2\pi}\sigma}\exp(\frac{-(\log(y)-\mu)^2}{2\sigma^2})$$

$\log(y) > \mu$, so $y > e^\mu = \theta$.

Let $\theta = e^\mu$, $\lambda = \sigma$,then $f(y) = \frac{2}{\sqrt{2\pi}\lambda y}\exp(\frac{-(\log(y)-\log(\theta))^2}{2\lambda^2}), y > \theta,$

$\Rightarrow Y \sim HPAR(\theta, \lambda)$.

Let $W = \log(Y)$, then $Y = r(W) = e^W$

$\Rightarrow g(w) = f_Y(r(w))|\frac{dr(w)}{dw}| = \frac{2}{\sqrt{2\pi}\lambda e^w}\exp[\frac{-(\log(e^w)-\log(\theta))^2}{2\lambda^2}]e^w$

$= \frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-(w-\mu)^2}{2\lambda^2})$,where $\mu = \log(\theta)$.

So $g(w) = \frac{2}{\sqrt{2\pi}\sigma}\exp(-\frac{(w-\mu)^2}{2\sigma^2})$, $w \geq \mu$, $\lambda = \sigma$

$\Rightarrow W \sim HN(\mu, \sigma)$.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $HPAR(\theta, \lambda)$, then the likelihood

$L(\theta, \lambda) = \prod_{i=1}^n f(y_i) = (\frac{2}{\sqrt{2\pi}\lambda})^n \frac{1}{\prod_{i=1}^n y_i}\exp(\sum_{i=1}^n(\frac{-(\log(y_i)-\log(\theta))^2}{2\lambda^2}))I(y_{(1)} \geq \theta)$,

and the log likelihood

$\log L = n\log(2) - n\log(\sqrt{2\pi}\lambda) + \sum_{i=1}^n -\log(y_i) + \sum_{i=1}^n \frac{-(\log y_i - \log\theta)^2}{2\lambda^2}$

$= c - n\log\lambda - \sum_{i=1}^n \log y_i - \sum_{i=1}^n \frac{(\log y_i - \log\theta)^2}{2\lambda^2}$.

In order to maximize $\log L$, we need to minimize $\sum_{i=1}^n(\log y_i - \log\theta)^2$ subject to the constraint $y_{(1)} \geq \theta$. This occurs when $\theta = y_{(1)}$. Hence MLE $\hat{\theta} = y_{(1)}$.

For this choice of θ,

$$\frac{d\log L}{d\lambda} = -\frac{n}{\lambda} - \sum_{i=1}^n \frac{-2(\log y_i - \log\theta)^2}{2\lambda^3}$$

$$= \frac{-n}{\lambda} + \frac{1}{\lambda^3}\sum_{i=1}^{n}(\log y_i - \log y_{(1)})^2 := 0$$

or

$$\lambda^2 = \frac{\sum_{i=1}^{n}(\log y_i - \log y_{(1)})^2}{n},$$

or

$$\hat{\lambda} = \sqrt{\frac{\sum_{i=1}^{n}(\log y_i - \log y_{(1)})^2}{n}}.$$

Notice that

$$\frac{d^2 \log L}{d\lambda^2}|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \frac{3}{\hat{\lambda}^4}\sum_{i=1}^{n}(\log y_i - \log y_{(1)})^2 = \frac{n}{\hat{\lambda}^2}[1-3] < 0$$

Hence $\hat{\lambda}$ is the MLE of λ.

Likelihood Based Confidence Intervals

Let $\sigma = \lambda, W = \log Y$ then $W \sim HN(\mu = \log(\theta), \sigma = \lambda)$, so (2.4) is a CI for $\sigma = \lambda$; that is, a large sample $100(1-\alpha)\%$CI for λ^2 if θ is unknown is

$$(\frac{n\hat{\lambda}^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{n\hat{\lambda}^2}{\chi^2_{n-1,\frac{\alpha}{2}}})$$

or

$$(\frac{\sum_{i=1}^{n}(\log y_i - \log y_{(1)})^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^{n}(\log y_i - \log y_{(1)})^2}{\chi^2_{n-1,\frac{\alpha}{2}}}). \tag{2.13}$$

If θ is known, then a large sample $100(1-\alpha)\%$CI for λ^2 is

$$(\frac{\sum_{i=1}^{n}(\log y_i - \log \theta)^2}{\chi^2_{n,1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^{n}(\log y_i - \log \theta)^2}{\chi^2_{n,\frac{\alpha}{2}}}). \tag{2.14}$$

Taking square roots of the endpoints gives a large sample $100(1-\alpha)\%$ CI for λ. A CI for $\mu = \log(\theta)$ is given by (2.7), that is, a large sample $100(1-\alpha)\%$ CI for $\log(\theta)$ is

$$(\log\hat{\theta} + \hat{\lambda}\log(\alpha)\Phi^{-1}(\frac{1}{2} + \frac{1}{2n})(1 + \frac{13}{n^2}), \log\hat{\theta})$$

so exponentiate the endpoints of (2.7) for a CI for θ :

$$(\hat{\theta}\exp(\hat{\lambda}\log(\alpha)\Phi^{-1}(\frac{1}{2}+\frac{1}{2n})(1+\frac{13}{n^2})),\hat{\theta})$$

or

$$(Y_{(1)}\exp(\hat{\lambda}\log(\alpha)\Phi^{-1}(\frac{1}{2}+\frac{1}{2n})(1+\frac{13}{n^2})),Y_{(1)}). \quad (2.15)$$

2.9 THE POWER DISTRIBUTION

If Y has a Power distribution, $Y \sim POW(\lambda)$, then the cdf $F_Y(y) = y^{\frac{1}{\lambda}},\ 0 \leq y \leq 1$, and the probability density function
$f(y) = \frac{1}{\lambda}y^{\frac{1}{\lambda}-1},\ 0 \leq y \leq 1, \lambda > 0.$

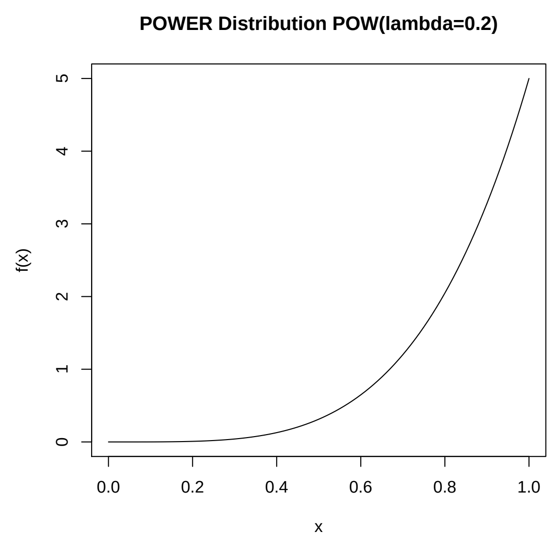

Figure 2.9. Plot of the pdf of the Power Distribution

Let $W = -\log(Y)$, then
$F_W(w) = P(W \leq w) = P(-\log(Y) \leq w) = P(\log(1/Y) \leq w) = P(\frac{1}{Y} \leq e^w) = P(Y \geq e^{-w}) = 1 - P(Y \leq e^{-w}) = 1 - F_Y(e^{-w}) = 1 - (\exp(-w))^{\frac{1}{\lambda}}) = 1 - \exp(\frac{-w}{\lambda}), w \geq 0$

$\Rightarrow W \sim EXP(\lambda)$.

If $W \sim EXP(\lambda)$, let $Y = e^{-W}$. Then

$F_Y(y) = P(Y \le y) = P(e^{-W} \le y) = P(-W \le \log(y)) = P(W \ge -\log(y))$

$= 1 - P(W \le -\log(y)) = 1 - F_W(-\log(y)) = 1 - (1 - \exp(-\frac{-\log(y)}{\lambda}))$

$= \exp(\frac{\log(y)}{\lambda}) = \exp(\log(y^{\frac{1}{\lambda}})) = y^{\frac{1}{\lambda}}, 0 \le y \le 1,\ \lambda > 0.$

$\Rightarrow Y \sim POW(\lambda)$.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $POW(\lambda)$, then the likelihood $L(\lambda) = \prod_{i=1}^n f(y_i) = (\frac{1}{\lambda})^n \prod_{i=1}^n y_i^{\frac{1}{\lambda}-1}$, and the log likelihood $\log L = -n \log \lambda + \sum_{i=1}^n (\frac{1}{\lambda} - 1) \log y_i$.

Hence

$$\frac{d \log L}{d\lambda} = \frac{-n}{\lambda} + \sum_{i=1}^n \frac{-1}{\lambda^2} \log y_i := 0$$

or

$$\frac{n}{\lambda} = \frac{-1}{\lambda^2} \sum_{i=1}^n \log y_i,$$

or

$$\hat{\lambda} = \frac{-\sum_{i=1}^n \log y_i}{n}.$$

Notice that

$$\frac{d^2 \log L}{d\lambda^2}|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} + \frac{2}{\hat{\lambda}^3} \sum_{i=1}^n 2 \log y_i = \frac{n}{\hat{\lambda}^2} + 2\frac{\sum_{i=1}^n \log y_i}{\hat{\lambda}^3} = \frac{n}{\hat{\lambda}^2} - 2n\frac{\sum_{i=1}^n \log y_i}{n\hat{\lambda}^3} = \frac{n}{\hat{\lambda}^2}[1-2] < 0.$$

Hence $\hat{\lambda}$ is the MLE of λ.

By (2.3), an exact $100(1-\alpha)\%$ confidence interval for λ is given by

$$\left(\frac{2T_n}{\chi^2_{2n,1-\alpha/2}}, \frac{2T_n}{\chi^2_{2n,\alpha/2}} \right), \tag{2.16}$$

where $T_n = \sum_{i=1}^n (W_i - 0) = -\sum_{i=1}^n \log Y_i$. Hence, a $100(1-\alpha)\%$ CI for λ is

$$\left(\frac{-2\sum_{i=1}^n \log Y_i}{\chi^2_{2n,1-\alpha/2}}, \frac{-2\sum_{i=1}^n \log Y_i}{\chi^2_{2n,\alpha/2}} \right). \tag{2.17}$$

2.10 THE HPOWER DISTRIBUTION

If $Y \sim HPOW(\lambda)$, then

$f(y) = \frac{2}{\sqrt{2\pi}} \frac{1}{\lambda y} \exp(\frac{-(\log(y))^2}{2\lambda^2}) I[0 \leq y \leq 1]$, $\lambda > 0$.

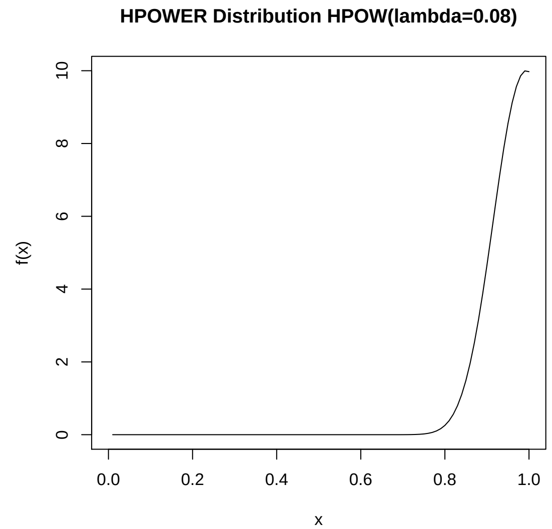

Figure 2.10. Plot of the pdf of the HPower Distribution

This distribution is unimodal with mode at $y = e^{-\lambda^2}$ and $f(e^{-\lambda^2}) = \sqrt{\frac{2}{\pi}} \frac{e^{\frac{\lambda^2}{2}}}{\lambda}$.

Proof:

$\log(f(y)) = c - \log y - \frac{(\log y)^2}{2\lambda^2}$.

So $\frac{d}{dy} \log f(y) = 0 + -\frac{1}{y} - \frac{1}{\lambda^2}(\log y)\frac{1}{y} = \frac{1}{y}[-1 - \frac{1}{\lambda^2}(\log(y)] := 0$

or

$1 = \frac{-1}{\lambda^2} \log(y)$, or $y = e^{-\lambda^2}$,

or

$\frac{d}{dy} f(y) = \frac{2}{\sqrt{2\pi}\lambda y} \exp(\frac{-(\log(y))^2}{2\lambda^2})(\frac{-2}{2\lambda^2} \log(y)\frac{1}{y}) + \exp(\frac{-(\log(y))^2}{2\lambda^2})(\frac{-2}{\sqrt{2\pi}\lambda y^2})$

$= \frac{-2}{\sqrt{2\pi}\lambda y^2} \exp(\frac{-(\log(y))^2}{2\lambda^2})[\frac{1}{\lambda^2} \log(y) + 1] := 0$, but the first two terms can't be zero, hence $\frac{1}{\lambda^2}\log(y)+1=0$ or $-\lambda^2=\log(y)$ or $y=e^{-\lambda^2}$. Notice that

$$\frac{d^2f(y)}{dy^2}|_{y=e^{-\lambda^2}} = \frac{-2}{\sqrt{2\pi}\lambda y^2}\exp(\frac{-(\log(y))^2}{2\lambda^2})(\frac{1}{\lambda^2 y}) + (\frac{1}{\lambda^2}\log(y)+1)[\frac{d}{dy}(\frac{-2}{\sqrt{2\pi}\lambda y^2}\exp(\frac{-(\log(y))^2}{2\lambda^2}))] < 0$$

since the first term is less than zero and the second term is zero because $(\frac{1}{\lambda^2}\log(y)+1)$ at $y=e^{-\lambda^2}$ is equal to $\frac{1}{\lambda^2}\log(e^{-\lambda^2})+1=-1+1=0$. Hence $y=e^{-\lambda^2}$ is a local maximum for $f(y)$. So the mode is at $y=e^{-\lambda^2}$.

If W has a half normal distribution, $W \sim HN(0,\lambda)$,

then $g_W(w)=\frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-(w-0)^2}{2\lambda^2}), w\geq 0,\ \lambda>0$.

Let $Y=e^{-W}$, then $W=s(Y)=-\log(Y)$

$\Rightarrow f(y)=g_W(s(y))|\frac{ds(y)}{dy}|=\frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-(-\log y)^2}{2\lambda^2})|-\frac{1}{y}|$

$=\frac{2}{\sqrt{2\pi}}\frac{1}{\lambda y}\exp(\frac{-(\log(y))^2}{2\lambda^2})I[0\leq y\leq 1]$

$\Rightarrow Y \sim HPOW(\lambda)$.

Let $W=-\log(Y)$,then $Y=r(W)=e^{-W}$

$\Rightarrow g(w)=f_Y(r(w))|\frac{dr(w)}{dw}|=\frac{2}{\sqrt{2\pi}\lambda}\frac{1}{e^{-w}}\exp[\frac{-(\log(e^{-w}))^2}{2\lambda^2}]I[0\leq e^{-w}\leq 1]|e^{-w}|$

$=\frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-w^2}{2\lambda^2})\,I[0\leq w<\infty]$

$\Rightarrow W \sim HN(0,\lambda)$.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $HPOW(\lambda)$, then the likelihood

$L(\lambda)=\prod_{i=1}^n f(y_i)=(\frac{2}{\sqrt{2\pi}\lambda})^n\frac{1}{\prod_{i=1}^n y_i}\exp\sum_{i=1}^n(\frac{-(\log^2(y_i))}{2\lambda^2})I(y_{(1)}\geq 0)I(y_{(n)}\leq 1)$,

and the log likelihood

$\log L=c-n\log\lambda-\sum_{i=1}^n\log y_i+\sum_{i=1}^n\frac{-\ [log(y_i)]^2}{2\lambda^2}$.

Hence

$$\frac{d\log L}{d\lambda}=\frac{-n}{\lambda}-\sum_{i=1}^n\frac{-2[\log y_i]^2}{2\lambda^3}$$

$$=\frac{-n}{\lambda}+\frac{\sum_{i=1}^n[\log y_i]^2}{\lambda^3}:=0$$

or

$$\frac{n}{\lambda} = \frac{\sum_{i=1}^{n}[\log y_i]^2}{\lambda^3},$$

or

$$\hat{\lambda} = \sqrt{\frac{\sum_{i=1}^{n}[\log y_i]^2}{n}}.$$

Notice that

$$\frac{d^2 \log L}{d\lambda^2}|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \frac{\sum_{i=1}^{n} 3[\log y_i]^2}{\hat{\lambda}^4} = \frac{n}{\hat{\lambda}^2}[1-3] < 0.$$

Hence $\hat{\lambda}$ is the MLE of λ.

Likelihood Based Confidence Intervals

Let $W = -\log Y$ then $W \sim HN(0, \lambda)$, then $W = 0 + \lambda X = \lambda X, X = |Z|$, where $Z \sim N(0,1)$. Since $\hat{\lambda} = \sqrt{\frac{\sum_{i=1}^{n}[\log Y_i]^2}{n}}$, then $n\hat{\lambda}^2 = \sum_{i=1}^{n}[\log Y_i]^2 = \sum_{i=1}^{n}[\log(e^{-W_i})]^2 = \sum_{i=1}^{n} W_i^2 = \lambda^2 \sum_{i=1}^{n} X_i^2 \Rightarrow \frac{n\hat{\lambda}^2}{\lambda^2} \sim \chi_n$. Hence a large sample $100(1-\alpha)\%$ CI for λ^2 is

$$\left(\frac{n\hat{\lambda}^2}{\chi^2_{n,1-\frac{\alpha}{2}}}, \frac{n\hat{\lambda}^2}{\chi^2_{n,\frac{\alpha}{2}}}\right) =$$

$$\left(\frac{\sum_{i=1}^{n}[\log(Y_i)]^2}{\chi^2_{n,1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^{n}[\log(Y_i)]^2}{\chi^2_{n,\frac{\alpha}{2}}}\right). \tag{2.18}$$

2.11 THE RAYLEIGH DISTRIBUTION

If Y has a Rayleigh distribution, $Y \sim R(\mu, \sigma)$, then the pdf of Y is

$$f(y) = \frac{y-\mu}{\sigma^2} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]$$

where $\sigma > 0$, μ is real, and $y \geq \mu$. The cdf of Y is

$$F(y) = 1 - \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right]$$

for $y \geq \mu$, and $F(y) = 0$, otherwise.

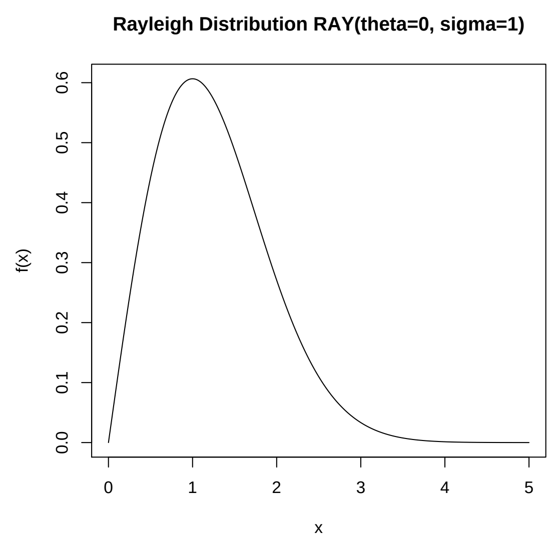

Figure 2.11. Plot of the pdf of the Rayleigh Distribution

Let $W = (Y-\mu)^2$, then $Y = r(W) = \mu + \sqrt{W}$, and

$F_W(w) = g_W(r(w))|\frac{dr(w)}{dw}| = \frac{(\mu+\sqrt{w})-\mu}{\sigma^2} \exp(-\frac{1}{2}(\frac{\mu+\sqrt{w}-\mu}{\sigma})^2)|\frac{1}{2\sqrt{w}}|$

$= \frac{1}{2\sigma^2} \exp(\frac{-w}{2\sigma^2}) \Rightarrow W \sim EXP(2\sigma^2)$.

If $W \sim EXP(2\sigma^2)$, let $Y = \sqrt{W} + \mu$. Then

$F_Y(y) = P(Y \leq y) = P((\sqrt{W} + \mu) \leq y) = P(\sqrt{W} \leq y - \mu) = P(W \leq (y-\mu)^2) =$

$1-\exp(\frac{-(y-\mu)^2}{2\sigma^2}), y \geq \mu,\ \sigma > 0$

$\Rightarrow Y \sim R(\mu, \sigma)$.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $R(\mu, \sigma)$, then the likelihood

$L(\sigma) = \prod_{i=1}^n f(y_i) = \prod_{i=1}^n (\frac{y_i-\mu}{\sigma^2}) \exp \sum_{i=1}^n -\frac{1}{2}(\frac{y_i-\mu}{\sigma})^2$,

and the log likelihood

$\log L(\sigma) = \sum_{i=1}^n \log(\frac{y_i-\mu}{\sigma^2}) + \sum_{i=1}^n -\frac{1}{2}(\frac{y_i-\mu}{\sigma})^2$.

Hence

$$\frac{d \log L}{d\sigma} = \sum_{i=1}^n \frac{\sigma^2}{y_i-\mu} \frac{(\mu-y_i)2\sigma}{\sigma^4} - \frac{1}{2}\sum_{i=1}^n 2(\frac{y_i-\mu}{\sigma})(\frac{\mu-y_i}{\sigma^2})$$
$$= \sum_{i=1}^n \frac{-2}{\sigma} + \sum_{i=1}^n \frac{(y_i-\mu)^2}{\sigma^3} := 0$$

or

$$\frac{2n}{\sigma} = \frac{\sum_{i=1}^n (y_i-\mu)^2}{\sigma^3},$$

or

$$\widehat{\sigma^2} = \frac{\sum_{i=1}^n (y_i-\mu)^2}{2n},$$

or

$$\widehat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (y_i-\mu)^2}{2n}}.$$

Notice that

$$\frac{d^2 \log L}{d\sigma^2}|_{\hat{\sigma}} = \sum_{i=1}^n \frac{2}{\hat{\sigma}^2} - \sum_{i=1}^n \frac{3(y_i-\mu)^2}{\hat{\sigma}^4} = \frac{2n}{\hat{\sigma}^2} - \frac{3(2n)}{\widehat{\sigma^4}} = \frac{2n}{\hat{\sigma}^2}[1-3] < 0.$$

Hence $\hat{\sigma}$ is the MLE of σ if μ is known.

The confidence interval for σ^2 when μ is known will be derived next. If μ is known, let $W = (Y - \mu)^2$, then $W \sim EXP(2\sigma^2)$. By (2.3), an approximate $100(1-\alpha)\%$ confidence interval for λ is given by

$$\left(\frac{2T_n}{\chi^2_{2n,1-\alpha/2}}, \frac{2T_n}{\chi^2_{2n,\alpha/2}} \right). \tag{2.19}$$

where $T_n = \sum_{i=1}^n (W_i - 0) = \sum_{i=1}^n (W_i) = \sum_{i=1}^n (Y_i - \mu)^2$. Set $\lambda = 2\sigma^2$ then, a $100(1-\alpha)\%$ CI for σ^2 is

$$\left(\frac{\sum_{i=1}^n (Y_i - \mu)^2}{\chi^2_{2n,1-\alpha/2}}, \frac{\sum_{i=1}^n (Y_i - \mu)^2}{\chi^2_{2n,\alpha/2}} \right). \tag{2.20}$$

Likelihood Based Confidence Intervals

If both μ and σ are unknown, then the $MLE(\hat{\mu}, \hat{\sigma})$ must be found before obtaining CIs. The log likelihood

$\log L(\mu, \sigma) = \sum_{i=1}^n \log(\frac{y_i - \mu}{\sigma^2}) + \sum_{i=1}^n -\frac{1}{2}(\frac{y_i - \mu}{\sigma})^2$

$= -2n \log(\sigma) + \sum_{i=1}^n \log(y_i - \mu) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2$.

Hence

$$\frac{d \log L}{d\sigma} = -\frac{2n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (y_i - \mu)^2 := 0$$

gives

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (y_i - \hat{\mu})^2}{2n}. \tag{2.21}$$

Also

$$\frac{d \log L}{d\mu} = -\sum_{i=1}^n (y_i - \mu)^{-1} + \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \mu) := 0$$

gives

$$\sum_{i=1}^n (y_i - \mu)^{-1} = \frac{1}{\sigma^2}\sum_{i=1}^n (y_i - \mu) = \frac{\sum_{i=1}^n y_i}{\sigma^2} - \frac{n\mu}{\sigma^2}$$

or

$$\mu = \frac{\sigma^2}{n}[\frac{\sum_{i=1}^n y_i}{\sigma^2} - \sum_{i=1}^n (y_i - \mu)^{-1}].$$

One way to find the MLE is by iteration using Newton's method, where starting values can be found using the method of moments. Newton's method is used to solve $\boldsymbol{g}(\boldsymbol{\theta}) = \boldsymbol{0}$ for $\boldsymbol{\theta}$, where the solution is called $\hat{\boldsymbol{\theta}}$, and uses

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - [\boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta}_k)}]^{-1}\boldsymbol{g}(\boldsymbol{\theta}_k)$$

where

$$\boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta})} = \begin{bmatrix} \frac{\partial}{\partial\theta_1} g_1(\boldsymbol{\theta}) & \dots & \frac{\partial}{\partial\theta_p} g_1(\boldsymbol{\theta}) \\ \vdots & & \vdots \\ \frac{\partial}{\partial\theta_1} g_p(\boldsymbol{\theta}) & \dots & \frac{\partial}{\partial\theta_p} g_p(\boldsymbol{\theta}) \end{bmatrix}.$$

If the MLE is the solution of the likelihood equations, then use $\boldsymbol{g}(\boldsymbol{\theta}) = (g_1(\boldsymbol{\theta}), ..., g_p(\boldsymbol{\theta}))^T$ where

$$g_i(\boldsymbol{\theta}) = \frac{\partial}{\partial\theta_i}\log(L(\boldsymbol{\theta})).$$

Let $\boldsymbol{\theta}_0$ be an initial estimator, such as the method of moments estimator of $\boldsymbol{\theta}$. Let $\boldsymbol{D} = \boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta})}$. Then

$$D_{ij} = \frac{\partial}{\partial\theta_j} g_i(\boldsymbol{\theta}) = \frac{\partial^2}{\partial\theta_i\partial\theta_j}\log(L(\boldsymbol{\theta})) = \sum_{k=1}^n \frac{\partial^2}{\partial\theta_i\partial\theta_j}\log(f(x_k|\boldsymbol{\theta})),$$

and

$$\frac{1}{n}D_{ij} = \frac{1}{n}\sum_{k=1}^n \frac{\partial^2}{\partial\theta_i\partial\theta_j}\log(f(X_k|\boldsymbol{\theta})) \xrightarrow{D} E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log(f(X|\boldsymbol{\theta}))\right].$$

Newton's method converges if the initial estimator is sufficiently close, but may diverge otherwise. Hence $\sqrt{n}$ consistent initial estimators are recommended. Newton's method is also popular because if the partial derivative and integration operations can be interchanged, then

$$\frac{1}{n}\boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta})} \xrightarrow{D} -\boldsymbol{I}(\boldsymbol{\theta}).$$

For example, the regularity conditions hold for a kP-REF. Then a 100 $(1-\alpha)\%$ large sample CI for θ_i is

$$\hat{\theta}_i \pm z_{1-\alpha/2}\sqrt{-\boldsymbol{D}_{ii}^{-1}}$$

where

$$\boldsymbol{D}^{-1} = \left[\boldsymbol{D}_{\boldsymbol{g}(\hat{\boldsymbol{\theta}})}\right]^{-1}.$$

This result follows because

$$\sqrt{-\boldsymbol{D}_{ii}^{-1}} \approx \sqrt{[\boldsymbol{I}^{-1}(\hat{\boldsymbol{\theta}})]_{ii}/n}.$$

Next, apply the above results to the Rayleigh (μ, σ) distribution (although no check has been made on whether the regularity conditions hold for the Rayleigh distribution which is not a 2P-REF).

$$L(\mu, \sigma) = \left(\prod \frac{y_i - \mu}{\sigma^2}\right) \exp\left[-\frac{1}{2\sigma^2}\sum (y_i - \mu)^2\right].$$

Notice that for fixed σ, $L(Y_{(1)}, \sigma) = 0$. Hence the MLE $\hat{\mu} < Y_{(1)}$. Now the log likelihood

$$\log(L(\mu, \sigma)) = \sum_{i=1}^{n} \log(y_i - \mu) - 2n\log(\sigma) - \frac{1}{2}\sum \frac{(y_i - \mu)^2}{\sigma^2}.$$

Hence $g_1(\mu, \sigma) =$

$$\frac{\partial}{\partial \mu}\log(L(\mu, \sigma)) = -\sum_{i=1}^{n} \frac{1}{y_i - \mu} + \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \mu) \stackrel{set}{=} 0,$$

and $g_2(\mu, \sigma) =$

$$\frac{\partial}{\partial \sigma}\log(L(\mu, \sigma)) = \frac{-2n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i - \mu)^2 \stackrel{set}{=} 0,$$

which has solution

$$\hat{\sigma}^2 = \frac{1}{2n}\sum_{i=1}^{n}(Y_i - \hat{\mu})^2.$$

To obtain initial estimators, let $\hat{\sigma}_M = \sqrt{S^2/0.429204}$ and $\hat{\mu}_M = \overline{Y} - 1.253314\hat{\sigma}_M$. These would be the method of moments estimators if $S_M^2 = \frac{n-1}{n}s^2$ was used instead of the sample variance S^2. Then use $\mu_0 = \min(\hat{\mu}_M, Y_{(1)} - \hat{\mu}_M)$ and $\sigma_0 = \sqrt{\sum(Y_i - \mu_0)^2/(2n)}$. Now $\boldsymbol{\theta} = (\mu, \sigma)^T$ and

$$\boldsymbol{D} \equiv \boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta})} = \begin{bmatrix} \frac{\partial}{\partial \mu} g_1(\boldsymbol{\theta}) & \frac{\partial}{\partial \sigma} g_1(\boldsymbol{\theta}) \\ \\ \frac{\partial}{\partial \mu} g_2(\boldsymbol{\theta}) & \frac{\partial}{\partial \sigma} g_2(\boldsymbol{\theta}) \end{bmatrix} =$$

$$\begin{bmatrix} -\sum_{i=1}^n \frac{1}{(y_i-\mu)^2} - \frac{n}{\sigma^2} & -\frac{2}{\sigma^3}\sum_{i=1}^n (y_i - \mu) \\ \\ -\frac{2}{\sigma^3}\sum_{i=1}^n (y_i - \mu) & \frac{2n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{i=1}^n (y_i - \mu)^2 \end{bmatrix}.$$

So

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \begin{bmatrix} -\sum_{i=1}^n \frac{1}{(y_i-\mu_k)^2} - \frac{n}{\sigma_k^2} & -\frac{2}{\sigma_k^3}\sum_{i=1}^n (y_i - \mu_k) \\ \\ -\frac{2}{\sigma_k^3}\sum_{i=1}^n (y_i - \mu_k) & \frac{2n}{\sigma_k^2} - \frac{3}{\sigma_k^4}\sum_{i=1}^n (y_i - \mu_k)^2 \end{bmatrix}^{-1} \boldsymbol{g}(\boldsymbol{\theta}_k)$$

where

$$\boldsymbol{g}(\boldsymbol{\theta}_k) = \begin{pmatrix} g_1(\boldsymbol{\theta}_k) \\ g_2(\boldsymbol{\theta}_k) \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \mu} \log L(\boldsymbol{\theta}_k) \\ \frac{\partial}{\partial \sigma} \log L(\boldsymbol{\theta}_k) \end{pmatrix} = \begin{pmatrix} -\sum_{i=1}^n \frac{1}{(y_i-\mu_k)} - \frac{1}{\sigma_k^2}\sum_{i=1}^n (y_i - \mu_k) \\ \\ \frac{-2n}{\sigma_k} + \frac{1}{\sigma_k^3}\sum_{i=1}^n (y_i - \mu_k)^2 \end{pmatrix}.$$

This formula was iterated for 100 steps resulting in $\boldsymbol{\theta}_{101} = (\mu_{101}, \sigma_{101})^T$. Then we took $\hat{\mu} = \min(\mu_{101}, 2Y_{(1)} - \mu_{101})$ and

$$\hat{\sigma} = \sqrt{\frac{1}{2n}\sum_{i=1}^n (Y_i - \hat{\mu})^2}.$$

Then $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma})^T$ and $\boldsymbol{D} \equiv \boldsymbol{D}_{\boldsymbol{g}(\hat{\boldsymbol{\theta}})}$. Then (assuming regularity conditions hold) a 100 $(1-\alpha)$% large sample CI for μ is

$$\hat{\mu} \pm z_{1-\alpha/2}\sqrt{-\boldsymbol{D}_{11}^{-1}} \tag{2.22}$$

and a 100 $(1-\alpha)$% large sample CI for σ is

$$\hat{\sigma} \pm z_{1-\alpha/2}\sqrt{-\boldsymbol{D}_{22}^{-1}}. \tag{2.23}$$

2.12 THE HRAYLEIGH DISTRIBUTION

If $Y \sim HRAY(\theta, \lambda)$, then $f(y) = \frac{4}{\sqrt{2\pi}\lambda}(y-\theta)\exp(\frac{-(y-\theta)^4}{2\lambda^2})$ for $\theta > 0,\ \lambda > 0,\ y \geq \theta$.

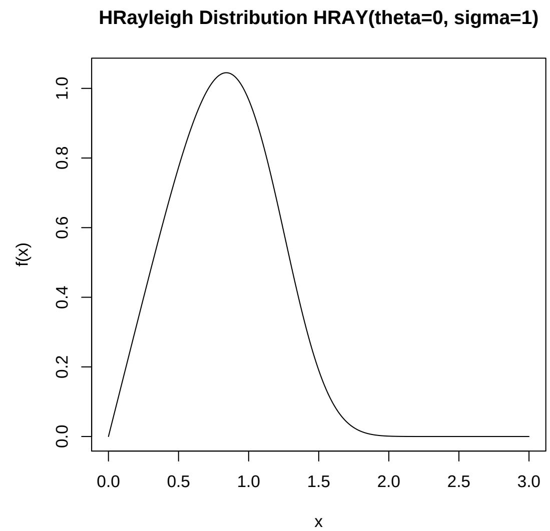

Figure 2.12. Plot of the pdf of the HRayleigh Distribution

If W has a half normal distribution, $W \sim HN(0, \lambda)$, then $g_W(w) = \frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-(w-0)^2}{2\lambda^2}), w \geq 0$. Let $Y = \sqrt{W} + \theta$, then $W = s(Y) = (Y-\theta)^2$

$\Rightarrow$

$$f(y) = g_W(s(y))|\frac{ds(y)}{dy}| = \frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-(y-\theta)^4}{2\lambda^2})|2(y-\theta)|$$
$$= \frac{4}{\sqrt{2\pi}\lambda}(y-\theta)\exp(\frac{-(y-\theta)^4}{2\lambda^2})$$

for $\theta > 0,\ \lambda > 0,\ y \geq \theta$,

$\Rightarrow Y \sim HRAY(\theta, \lambda)$.

Let $W = (Y-\theta)^2$, then $Y = r(W) = \sqrt{W} + \theta$

$\Rightarrow g(w) = f(r(w))|\frac{dr(w)}{dw}| = \frac{4}{\sqrt{2\pi}\lambda}(\sqrt{w} + \theta - \theta)\exp[\frac{-(\sqrt{w}+\theta-\theta)^4}{2\lambda^2}]|\frac{1}{2\sqrt{w}}|$

$= \frac{4}{\sqrt{2\pi}\lambda}\sqrt{w}\exp(\frac{-w^2}{2\lambda^2})\frac{1}{2\sqrt{w}}$

$= \frac{2}{\sqrt{2\pi}\lambda} \exp(\frac{-w^2}{2\lambda^2}),\ w \geq 0$

$\Rightarrow W \sim HN(0, \lambda).$

2.13 THE SMALLEST EXTREME VALUE DISTRIBUTION

If Y has a smallest extreme value distribution, $Y \sim SEV(\theta, \sigma)$, then the pdf of Y is

$$f(y) = \frac{1}{\sigma} \exp(\frac{y-\theta}{\sigma}) \exp[-\exp(\frac{y-\theta}{\sigma})]$$

where y and θ are real and $\sigma > 0$. This distribution is a location scale family. The cdf of Y is

$$F(y) = 1 - \exp[-\exp(\frac{y-\theta}{\sigma})].$$

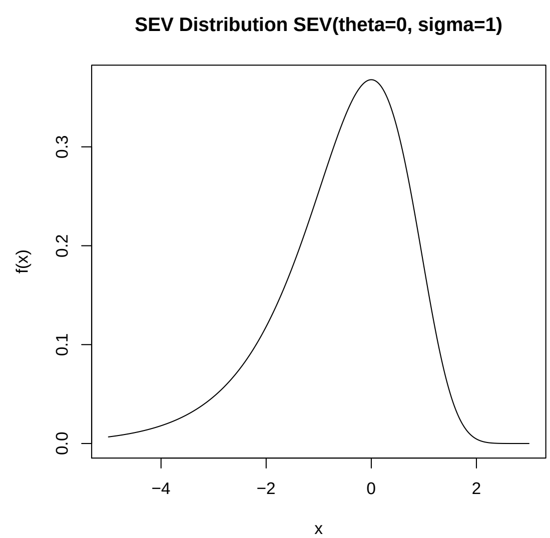

Figure 2.13. Plot of the pdf of the Smallest Extreme Value Distribution

Let $W = \exp(\frac{(Y-\theta)}{\sigma})$. Then $F_W(w) = P(W \leq w) = P(\exp(\frac{Y-\theta}{\sigma}) \leq w) = P(\frac{Y-\theta}{\sigma} \leq \log w) = P(Y - \theta \leq \sigma \log w) = P(Y \leq \theta + \sigma \log w) = 1 - \exp(-\exp(\frac{(\theta+\sigma \log w)-\theta}{\sigma})) = 1 - \exp(-w)$

$\Rightarrow W \sim EXP(1).$

If $W \sim EXP(1)$, let $Y = \sigma \log W + \theta$. Then

$F_Y(y) = P(Y \le y) = P(\sigma \log W + \theta \le y) = P(\log W \le \frac{y-\theta}{\sigma}) = P(W \le \exp(\frac{y-\theta}{\sigma})) = 1 - \exp(-\exp(\frac{(y-\theta)}{\sigma})), \sigma \ge 0, -\infty < y < \infty.$

$\Rightarrow Y \sim SEV(\theta, \sigma).$

2.14 THE HSEV DISTRIBUTION

If $Y \sim HSEV(\theta, \lambda)$, then the pdf of Y is

$$f(y) = \frac{2}{\sqrt{2\pi}\lambda} \exp(\frac{y-\theta}{\lambda}) \exp(-\frac{1}{2}[\exp(\frac{y-\theta}{\lambda})]^2), y \in R, \theta \in R, \lambda > 0$$

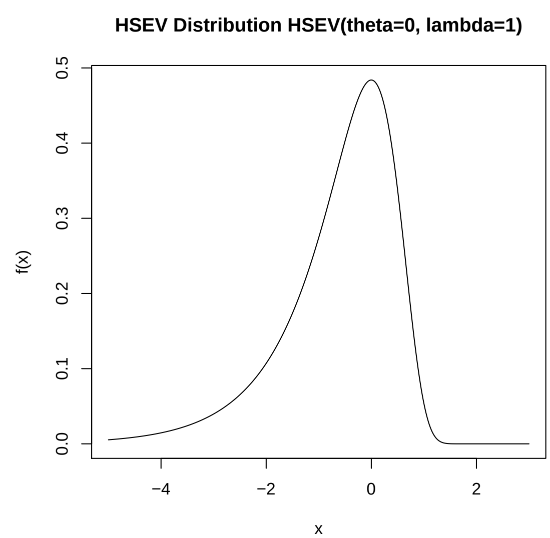

Figure 2.14. Plot of the pdf of the HSEV Distribution

If W has a half normal distribution, $W \sim HN(0,1)$, then $g_W(w) = \frac{2}{\sqrt{2\pi}} \exp(\frac{-w^2}{2})$ for $w \ge 0$. Let $Y = \lambda \log(W) + \theta$, then $W = s(Y) = \exp(\frac{Y-\theta}{\lambda})$

$\Rightarrow f(y) = g_W(s(y))|\frac{ds(y)}{dy}| = \frac{2}{\sqrt{2\pi}} \exp(\frac{-[\exp(\frac{y-\theta}{\lambda})]^2}{2})|\exp(\frac{y-\theta}{\lambda})\frac{1}{\lambda}|$

$= \frac{2}{\sqrt{2\pi}\lambda} \exp(\frac{(y-\theta)}{\lambda}) \exp(-\frac{1}{2}[\exp(\frac{y-\theta}{\lambda})]^2), y \in R, \theta \in R, \lambda > 0$

$\Rightarrow Y \sim HSEV(\theta, \lambda).$

Let $W = \exp(\frac{Y-\theta}{\lambda})$,then $Y = \lambda\log(W) + \theta \Rightarrow g(w) = f(r(w))|\frac{dr(w)}{dw}|$

$= \frac{2}{\sqrt{2\pi}\lambda}\exp[\frac{\lambda\log(w)+\theta-\theta}{\lambda}]\exp[-\frac{1}{2}(\exp(\frac{\lambda\log(w)+\theta-\theta}{\lambda}))^2]|\frac{\lambda}{w}|$

$= \frac{2}{\sqrt{2\pi}}w\exp(\frac{-w^2}{2})\frac{1}{w} = \frac{2}{\sqrt{2\pi}}\exp(\frac{-w^2}{2}),\ w \geq 0$

$\Rightarrow W \sim HN(0,1)$.

2.15 THE TRUNCATED EXTREME VALUE DISTRIBUTION

If Y has a Truncated Extreme Value distribution, $Y \sim TEV(\lambda)$, then $F_Y(y) = 1 - \exp(\frac{-(e^y-1)}{\lambda})$, $y > 0$, and$f(y) = \frac{1}{\lambda}\exp(y - \frac{e^y-1}{\lambda}), y > 0,\ \lambda > 0$.

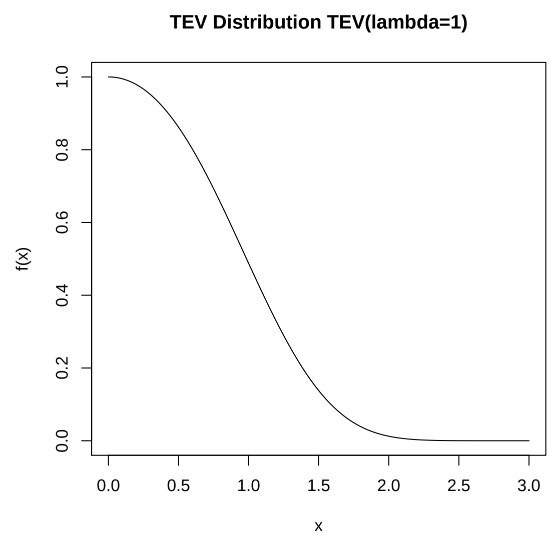

Figure 2.15. Plot of the pdf of the Truncated Extreme Value Distribution

Let $W = e^Y - 1$, then $F_W(w) = P(W \leq w) = P(e^Y - 1 \leq w) = P(e^Y \leq 1+w)$

$= P(Y \leq \log(1+w)) = 1 - (\exp(-\frac{e^{log(1+w)}-1}{\lambda})) = 1 - (\exp(-\frac{(1+w)-1}{\lambda}))$

$= 1 - e^{-w/\lambda}$

$\Rightarrow W \sim EXP(\lambda)$.

If $W \sim EXP(\lambda)$, let $Y = \log(1+w)$. Then

$F_Y(y) = P(Y \leq y) = P(\log(1+W) \leq y) = P(1+W \leq e^y) = P(W \leq e^Y - 1)$

$= 1 - \exp(-\frac{e^y-1}{\lambda}), y > 0$

$\Rightarrow Y \sim TEV(\lambda)$.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $TEV(\lambda)$, then the likelihood

$L(\lambda) = \prod_{i=1}^n f(y_i) = (\frac{1}{\lambda})^n \exp \sum_{i=1}^n (y_i - \frac{e_i^y - 1}{\lambda})$,

and the log likelihood

$\log L = -n \log \lambda + \sum_{i=1}^n (y_i - \frac{e_i^y-1}{\lambda})$.

Hence

$$\frac{d \log L}{d\lambda} = \frac{-n}{\lambda} + \sum_{i=1}^n \frac{e_i^y - 1}{\lambda^2} := 0$$

or

$$\hat{\lambda} = \frac{\sum_{i=1}^n (e_i^y - 1)}{n}.$$

Notice that

$$\frac{d^2 \log L}{d\lambda^2}|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \sum_{i=1}^n \frac{2(e_i^y - 1)}{\hat{\lambda}^3} = \frac{n}{\hat{\lambda}^2} - \frac{2n}{\hat{\lambda}^3}\hat{\lambda} = \frac{n}{\hat{\lambda}^2}[1-2] < 0.$$

Hence $\hat{\lambda}$ is the MLE of λ.

Likelihood Based Confidence Intervals

Let $W = e^Y - 1$, then $W \sim EXP(\lambda)$. By (2.3), an exact $100(1-\alpha)\%$ confidence interval for λ is given by

$$\left(\frac{2T_n}{\chi^2_{2n,1-\alpha/2}}, \frac{2T_n}{\chi^2_{2n,\alpha/2}} \right). \tag{2.24}$$

where $T_n = \sum_{i=1}^n (W_i - 0) = \sum_{i=1}^n W_i = \sum_{i=1}^n (e^{Y_i} - 1)$

Hence, a $100(1-\alpha)\%$ CI for λ is

$$\left(\frac{2\sum_{i=1}^n (e^{Y_i} - 1)}{\chi^2_{2n,1-\alpha/2}}, \frac{2\sum_{i=1}^n (e^{Y_i} - 1)}{\chi^2_{2n,\alpha/2}} \right). \tag{2.25}$$

2.16 THE HTEV DISTRIBUTION

If $Y \sim HTEV(\lambda)$, then $f(y) = \frac{2}{\sqrt{2\pi}\lambda} e^y \exp(\frac{-(e^y-1)^2}{2\lambda^2}) I(y \geq 0),\ \lambda > 0.$

This distribution is unimodal and the mode at y_m is found by solving the following equation:

$\lambda^2 = e^{y_m}(e^{y_m} - 1).$

Proof:

$\log(f(y)) = \log(\frac{2}{\sqrt{2\pi}\lambda}) + y - \frac{(e^y-1)^2}{2\lambda^2}$

So $\frac{d}{dy} \log f(y) = 0 + 1 - (2(e^y - 1)(\frac{e^y}{2\lambda^2}) := 0$

or

$1 = \frac{2e^y(e^y-1)}{2\lambda^2} \Rightarrow \lambda^2 = e^y(e^y - 1).$

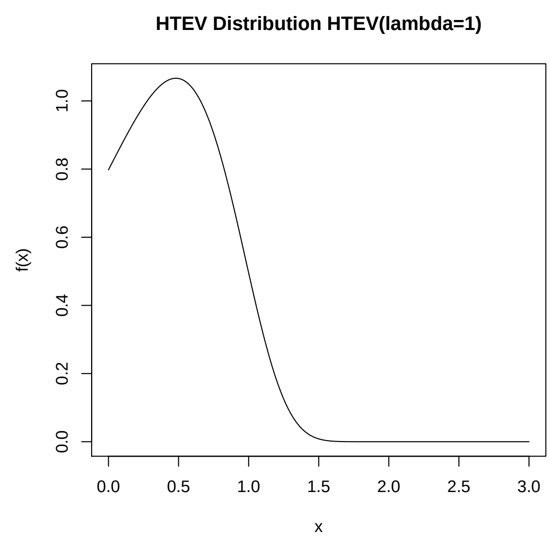

Figure 2.16. Plot of the pdf of the HTEV Distribution

If W has a half normal distribution, $W \sim HN(0, \lambda)$, then $g_W(w) = \frac{2}{\sqrt{2\pi}\lambda} \exp(\frac{-(w-0)^2}{2\lambda^2})$. for $w \geq 0$. Let $Y = \log(W + 1)$, then $W = s(Y) = e^Y - 1$

$\Rightarrow f(y) = g_W(s(y))|\frac{ds(y)}{dy}| = \frac{2}{\sqrt{2\pi}\lambda} \exp(\frac{-(e^y-1)^2}{2\lambda^2})|e^y|$

$= \frac{2}{\sqrt{2\pi}\lambda} e^y \exp(\frac{-(e^y-1)^2}{2\lambda^2}) I(y \geq 0)$

$\Rightarrow Y \sim HTEV(\lambda)$.

Let $W = e^Y - 1$, then $Y = r(W) = \log(W+1)$

$\Rightarrow g(w) = f(r(w))|\frac{dr(w)}{dw}|$

$= \frac{2}{\sqrt{2\pi}\lambda} e^{\log(w+1)} \exp[\frac{-(e^{log(w+1)}-1)^2}{2\lambda^2}] \frac{1}{w+1} I[\log(w+1) \geq 0]$

$= \frac{2}{\sqrt{2\pi}\lambda}(w+1)\exp(\frac{-w^2}{2\lambda^2})\frac{1}{w+1}\, I[w \geq 0] = \frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-w^2}{2\lambda^2})$

$\Rightarrow W \sim HN(0, \lambda)$.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $HTEV(\lambda)$, then the likelihood

$L(\lambda) = \prod_{i=1}^n f(y_i) = (\frac{2}{\sqrt{2\pi}\lambda})^n \exp(\sum_{i=1}^n y_i) \exp(\sum_{i=1}^n \frac{-(e^{y_i}-1)^2}{2\lambda^2}) I(y_{(1)} \geq 0)$,

and the log likelihood

$\log L = n\log(2) - n\log(\sqrt{2\pi}) - n\log\lambda + \sum_{i=1}^n y_i - \sum_{i=1}^n \frac{(e^{y_i}-1)^2}{2\lambda^2}$.

Hence

$$\frac{d\log L}{d\lambda} = \frac{-n}{\lambda} - \sum_{i=1}^n \frac{-2(e^{y_i}-1)^2}{2\lambda^3} = \frac{-n}{\lambda} + \frac{\sum_{i=1}^n (e^{y_i}-1)^2}{\lambda^3} := 0$$

or

$$\lambda^2 = \frac{\sum_{i=1}^n (e^{y_i}-1)^2}{n},$$

or

$$\hat{\lambda} = \sqrt{\frac{\sum_{i=1}^n (e^{y_i}-1)^2}{n}}.$$

Notice that

$$\frac{d^2\log L}{d\lambda^2}|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \frac{\sum_{i=1}^n 3(e^{y_i}-1)^2}{\hat{\lambda}^4} = \frac{n}{\hat{\lambda}^2}[1-3] < 0.$$

Hence $\hat{\lambda}$ is the MLE of λ.

Likelihood Based Confidence Intervals

Let $W = (e^Y - 1)$ then $W \sim HN(0, \lambda)$, and $W = 0 + \lambda X = \lambda X, X = |Z|$,

$Z \sim \mathrm{N}(0,1)$, then

Since $\hat{\lambda} = \sqrt{\frac{\sum_{i=1}^{n}(e^{y_i}-1)^2}{n}}$, we have $n\hat{\lambda}^2 = \sum_{i=1}^{n}(e^{y_i}-1)^2 = \sum_{i=1}^{n} w_i^2$

$= \lambda^2 \sum_{i=1}^{n} x_i^2$

$\Rightarrow \frac{n\hat{\lambda}^2}{\lambda^2} \sim \chi_n^2$.

Hence a large sample $100(1-\alpha)\%$ CI for λ^2 is

$$\left(\frac{n\hat{\lambda}^2}{\chi^2_{n,1-\frac{\alpha}{2}}}, \frac{n\hat{\lambda}^2}{\chi^2_{n,\frac{\alpha}{2}}}\right)$$

or

$$\left(\frac{\sum_{i=1}^{n}(e^{Y_i}-1)^2}{\chi^2_{n,1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^{n}(e^{Y_i}-1)^2}{\chi^2_{n,\frac{\alpha}{2}}}\right). \tag{2.26}$$

2.17 THE WEIBULL DISTRIBUTION

If Y has a Weibull distribution, $Y \sim W(\phi, \lambda)$, then $F_Y(y) = 1 - \exp(\frac{-y^{\phi}}{\lambda}), y > 0$, and

$f(y) = \frac{\phi}{\lambda} y^{\phi-1} \exp(\frac{-y^{\phi}}{\lambda})$, $y > 0,\ \phi \geq 0,\ \lambda > 0.$

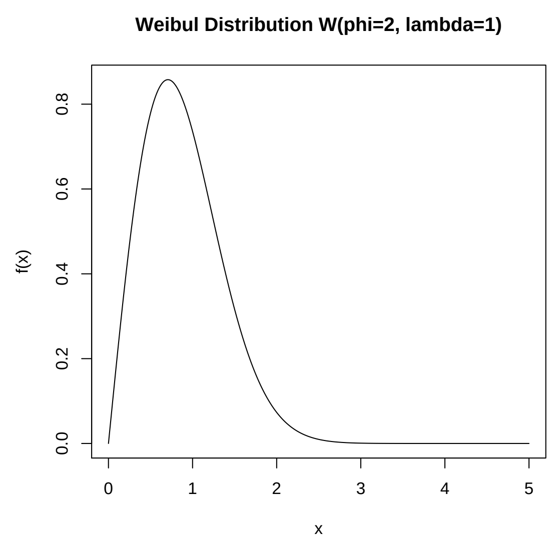

Figure 2.17. Plot of the pdf of the Weibull Distribution

Let $W = Y^{\phi}$, then

$F_W(w) = P(W \leq w) = P(Y^{\phi} \leq w) = P(Y \leq w^{\frac{1}{\phi}}) = F_Y(w^{\frac{1}{\phi}})$

$= 1 - \exp(-(w^{\frac{1}{\phi}})^{\phi}/\lambda) = 1 - \exp(\frac{-w}{\lambda})$

$\Rightarrow W \sim EXP(\lambda)$.

If $W \sim EXP(\lambda)$, let $Y = W^{\frac{1}{\phi}}$. Then

$F_Y(y) = P(Y \leq y) = P(W^{\frac{1}{\phi}} \leq y) = P(W \leq y^{\phi}) = F_W(y^{\phi}) = 1 - e^{\frac{-y^{\phi}}{\lambda}}$

$\Rightarrow f(y) = \frac{\phi}{\lambda} y^{\phi-1} \exp^{\frac{-y^{\phi}}{\lambda}}$

$\Rightarrow Y \sim W(\phi, \lambda)$.

Theorem 2.17.1: the Multivariate Central Limit Theorem (MCLT). If $\boldsymbol{X}_1, ..., \boldsymbol{X}_n$ are iid $k \times 1$ random vectors with $E(\boldsymbol{X}) = \boldsymbol{\mu}$ and $\text{Cov}(\boldsymbol{X}) = \boldsymbol{\Sigma}$, then

$$\sqrt{n}(\overline{\boldsymbol{X}} - \boldsymbol{\mu}) \xrightarrow{D} N_k(\boldsymbol{0}, \boldsymbol{\Sigma})$$

where the sample mean

$$\overline{\boldsymbol{X}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_i$$

see [12, 13].

Theorem 2.17.2: the Multivariate Delta Method. If

$$\sqrt{n}(\boldsymbol{T}_n - \boldsymbol{\theta}) \xrightarrow{D} N_k(\boldsymbol{0}, \boldsymbol{\Sigma}),$$

then

$$\sqrt{n}(\boldsymbol{g}(\boldsymbol{T}_n) - \boldsymbol{g}(\boldsymbol{\theta})) \xrightarrow{D} N_d(\boldsymbol{0}, \boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta})} \boldsymbol{\Sigma} \boldsymbol{D}^T_{\boldsymbol{g}(\boldsymbol{\theta})})$$

where the $d \times k$ Jacobian matrix of partial derivatives

$$\boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta})} = \begin{bmatrix} \frac{\partial}{\partial \theta_1} g_1(\boldsymbol{\theta}) & \dots & \frac{\partial}{\partial \theta_k} g_1(\boldsymbol{\theta}) \\ \vdots & & \vdots \\ \frac{\partial}{\partial \theta_1} g_d(\boldsymbol{\theta}) & \dots & \frac{\partial}{\partial \theta_k} g_d(\boldsymbol{\theta}) \end{bmatrix}.$$

Here the mapping $\boldsymbol{g} : \Re^k \rightarrow \Re^d$ needs to be differentiable in a neighborhood of $\boldsymbol{\theta} \in \Re^k$.

If $Y_1, ..., Y_n$ are iid Weibull (ϕ, λ), then the MLE $(\hat{\phi}, \hat{\lambda})$ must be found before obtaining CIs. The likelihood

$$L(\phi, \lambda) = \frac{\phi^n}{\lambda^n} \prod_{i=1}^{n} y_i^{\phi-1} \frac{1}{\lambda^n} \exp\left[\frac{-1}{\lambda} \sum y_i^{\phi}\right],$$

and the log likelihood

$$\log(L(\phi, \lambda)) = n\log(\phi) - n\log(\lambda) + (\phi - 1)\sum_{i=1}^{n} \log(y_i) - \frac{1}{\lambda}\sum y_i^{\phi}.$$

Hence

$$\frac{\partial}{\partial \lambda} \log(L(\phi, \lambda)) = \frac{-n}{\lambda} + \frac{\sum y_i^{\phi}}{\lambda^2} \stackrel{set}{=} 0,$$

or $\sum y_i^{\phi} = n\lambda$, or

$$\hat{\lambda} = \frac{\sum y_i^{\hat{\phi}}}{n}.$$

Notice that

$$\frac{\partial}{\partial \phi} \log(L(\phi, \lambda)) = \frac{n}{\phi} + \sum_{i=1}^{n} \log(y_i) - \frac{1}{\lambda}\sum y_i^{\phi} \log(y_i) \stackrel{set}{=} 0,$$

so

$$n + \phi[\sum_{i=1}^{n} \log(y_i) - \frac{1}{\lambda}\sum y_i^{\phi} \log(y_i)] = 0,$$

or

$$\hat{\phi} = \frac{n}{\frac{1}{\hat{\lambda}}\sum y_i^{\hat{\phi}} \log(y_i) - \sum_{i=1}^{n} \log(y_i)}.$$

One way to find the MLE is to use iteration [5]

$$\hat{\lambda}_k = \frac{\sum y_i^{\hat{\phi}_{k-1}}}{n}$$

and

$$\hat{\phi}_k = \frac{n}{\frac{1}{\hat{\lambda}_k}\sum y_i^{\hat{\phi}_{k-1}} \log(y_i) - \sum_{i=1}^{n} \log(y_i)}.$$

Since $W = \log(Y) \sim SEV(\theta = \log(\lambda^{1/\phi}), \sigma = 1/\phi)$, Olive [14] gave the following robust estimators for σ and ϕ:

$$\hat{\sigma}_R = MAD(W_1, ..., W_n)/0.767049$$

and

$$\hat{\theta}_R = MED(W_1, ..., W_n) - \log(\log(2))\hat{\sigma}_R.$$

Then $\hat{\phi}_0 = 1/\hat{\sigma}_R$ and $\hat{\lambda}_0 = \exp(\hat{\theta}_R/\hat{\sigma}_R)$. The iteration might be run until both $|\hat{\phi}_k - \hat{\phi}_{k-1}| < 10^{-6}$ and $|\hat{\lambda}_k - \hat{\lambda}_{k-1}| < 10^{-6}$. Then take $(\hat{\phi}, \hat{\lambda}) = (\hat{\phi}_k, \hat{\lambda}_k)$. If $\mu = \lambda^{1/\phi}$ so $\mu^{\phi} = \lambda$, and $Y \sim Weibull(\phi, \mu)$ then the Weibull pdf

$$f(y) = \frac{\phi}{\mu}\left(\frac{y}{\mu}\right)^{\phi-1} \exp\left[-\left(\frac{y}{\mu}\right)^{\phi}\right].$$

Let $(\hat{\mu}, \hat{\phi})$ be the MLE of (μ, ϕ). According to Bain (1978, p. 215) [2],

$$\sqrt{n}\left(\begin{pmatrix} \hat{\mu} \\ \hat{\phi} \end{pmatrix} - \begin{pmatrix} \mu \\ \phi \end{pmatrix} \right) \xrightarrow{D} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1.109\frac{\mu^2}{\phi^2} & 0.257\mu \\ 0.257\mu & 0.608\phi^2 \end{pmatrix} \right).$$

Let column vectors $\boldsymbol{\theta} = (\mu \;\; \phi)^T$ and $\boldsymbol{\eta} = (\lambda \;\; \phi)^T$. Then

$$\boldsymbol{\eta} = \boldsymbol{g}(\boldsymbol{\theta}) = \begin{pmatrix} \lambda \\ \phi \end{pmatrix} = \begin{pmatrix} \mu^{\phi} \\ \phi \end{pmatrix} = \begin{pmatrix} g_1(\boldsymbol{\theta}) \\ g_2(\boldsymbol{\theta}) \end{pmatrix}.$$

So

$$\boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta})} = \begin{bmatrix} \frac{\partial}{\partial\theta_1}g_1(\boldsymbol{\theta}) & \frac{\partial}{\partial\theta_2}g_1(\boldsymbol{\theta}) \\ \frac{\partial}{\partial\theta_1}g_2(\boldsymbol{\theta}) & \frac{\partial}{\partial\theta_2}g_2(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial\mu}\mu^{\phi} & \frac{\partial}{\partial\phi}\mu^{\phi} \\ \frac{\partial}{\partial\mu}\phi & \frac{\partial}{\partial\phi}\phi \end{bmatrix} = \begin{bmatrix} \phi\mu^{\phi-1} & \mu^{\phi}\log(\mu) \\ 0 & 1 \end{bmatrix}.$$

Thus by the multivariate delta method (Theorem 2.17.2),

$$\sqrt{n}\left(\begin{pmatrix} \hat{\lambda} \\ \hat{\phi} \end{pmatrix} - \begin{pmatrix} \lambda \\ \phi \end{pmatrix} \right) \xrightarrow{D} N_2(\boldsymbol{0}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\Sigma} = \boldsymbol{I}^{-1}(\boldsymbol{\eta}) = [\boldsymbol{I}(\boldsymbol{g}(\boldsymbol{\theta}))]^{-1} = \boldsymbol{D}_{\boldsymbol{g}(\boldsymbol{\theta})}\boldsymbol{I}^{-1}(\boldsymbol{\theta})\boldsymbol{D}^T_{\boldsymbol{g}(\boldsymbol{\theta})}$$

$$= \begin{bmatrix} \phi\mu^{\phi-1} & \mu^{\phi}\log\mu \\ & \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1.109\frac{\mu^2}{\phi^2} & .257\mu \\ & \\ .257\mu & .608\phi^2 \end{bmatrix} \begin{bmatrix} \phi\mu^{\phi-1} & 0 \\ & \\ \mu^{\phi}\log\mu & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \phi\mu^{\phi-1} & \mu^{\phi}\log\mu \\ & \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1.109\frac{\mu^{\phi+1}}{\phi} + .257\mu^{\phi+1}\log\mu & .257\mu \\ & \\ .257\phi\mu^{\phi} + .608\phi^2\mu^{\phi}\log\mu & .608\phi^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1.109\mu^{2\phi} + .514\phi\mu^{2\phi}\log\mu + .608\phi^2\mu^{2\phi}\log^2\mu & .257\phi\mu^{\phi} + .608\phi^2\mu^{\phi}\log\mu \\ & \\ .257\phi\mu^{\phi} + .608\phi^2\mu^{\phi}\log\mu & .608\phi^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1.109\lambda^2 + .514\phi\lambda^2\frac{\log\lambda}{\phi} + .608\phi^2\lambda^2\frac{\log^2\lambda}{\phi^2} & .257\phi\lambda + .608\phi\lambda\log\lambda \\ & \\ .257\phi\lambda + .608\phi\lambda\log\lambda & .608\phi^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1.109\lambda^2(1 + .4635\log\lambda + .5482(\log\lambda)^2) & .257\phi\lambda + .608\phi\lambda\log\lambda \\ & \\ .257\phi\lambda + .608\phi\lambda\log\lambda & .608\phi^2 \end{bmatrix}.$$

Hence the asymptotic variances of $\hat{\phi}$ and $\hat{\lambda}$ are given by

$AV(\hat{\phi}) = .608\hat{\phi}^2$ and

$AV(\hat{\lambda}) = 1.109\hat{\lambda}^2(1 + .4635\log\hat{\lambda} + .5482(\log\hat{\lambda})^2)$.

Hence

$\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{D} N(0, .608\phi^2)$.

Thus $1-\alpha \approx P(-z_{1-\alpha/2}\sqrt{0.608}\hat{\phi} < \sqrt{n}(\hat{\phi}-\phi) < z_{1-\alpha/2}\sqrt{0.608}\hat{\phi})$ and a large sample $100(1-\alpha)\%$ CI for ϕ is

$$\hat{\phi} \pm z_{1-\alpha/2}\ \hat{\phi}\ \sqrt{0.608/n}. \tag{2.27}$$

Similarly,

$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{D} N(0, 1.109\lambda^2(1 + .4635\log\lambda + .5482(\log\lambda)^2))$,

and a large sample $100(1-\alpha)\%$ CI for λ is

$$\hat{\lambda} \pm \frac{z_{1-\alpha/2}}{\sqrt{n}}\ \sqrt{1.109\hat{\lambda}^2[1 + 0.4635\log(\hat{\lambda}) + 0.5824(\log(\hat{\lambda}))^2]}. \tag{2.28}$$

In simulations, for small n the number of iterations for the MLE to converge could be in the thousands, and the coverage of the large sample CIs is poor for $n < 50$.

2.18 THE HWEIBULL DISTRIBUTION

If $Y \sim HW(\phi, \lambda)$, then $f(y) = \frac{2}{\sqrt{2\pi}\lambda}\phi y^{\phi-1}\exp(\frac{-y^{2\phi}}{2\lambda^2}), y > 0\, \lambda > 0$ and $\phi > 0$.

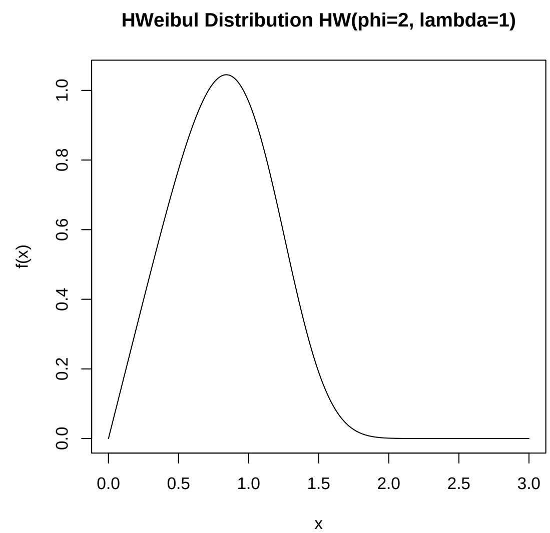

Figure 2.18. Plot of the pdf of the HWeibull Distribution

If W has a half normal distribution, $W \sim HN(0, \lambda)$, then

$g_W(w) = \frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-(w-0)^2}{2\lambda^2}), w \geq 0$. Let $Y = W^{\frac{1}{\phi}}$, then $W = s(Y) = Y^{\phi}$

$\Rightarrow f(y) = g_W(s(y))|\frac{ds(y)}{dy}| = \frac{2}{\sqrt{2\pi}\lambda}\exp(\frac{-y^{2\phi}}{2\lambda^2})|\phi y^{\phi-1}|$

$= \frac{2}{\sqrt{2\pi}\lambda}\phi y^{\phi-1}\exp(\frac{-y^{2\phi}}{2\lambda^2}), y > 0,$

$\Rightarrow Y \sim HW(\phi, \lambda)$.

Let $W = Y^{\phi}$, then $Y = r(W) = W^{\frac{1}{\phi}}$

$\Rightarrow g(w) = f(r(w))|\frac{dr(w)}{dw}|$

$= \frac{2}{\sqrt{2\pi}\lambda}\phi(w^{\frac{1}{\phi}})^{\phi-1}\exp[\frac{-(w^{\frac{1}{\phi}})^{2\phi}}{2\lambda^2}]\frac{1}{\phi}w^{\frac{1-\phi}{\phi}}$

$= \frac{2}{\sqrt{2\pi}\lambda} \exp(\frac{-w^2}{2\lambda^2})\, I[w \geq 0]$

$\Rightarrow W \sim HN(0, \lambda)$.

Let $Y_1, Y_2, \ldots, Y_n$ be iid $HW(\phi, \lambda)$,and if ϕ is known, then the likelihood

$L(\lambda) = \prod_{i=1}^n f(y_i) = (\frac{2}{\sqrt{2\pi}\lambda})^n \phi^n \prod_{i=1}^n y_i^{\phi-1} \exp(\sum_{i=1}^n \frac{-y_i^{2\phi}}{2\lambda^2})$,
and the log likelihood
$\log L = n\log(2) - n\log(\sqrt{2\pi}) - n\log\lambda + n\log\phi + \sum_{i=1}^n(\phi-1)\log y_i - \sum_{i=1}^n \frac{y_i^{2\phi}}{2\lambda^2}$.

Hence

$$\frac{d\log L}{d\lambda} = \frac{-n}{\lambda} - \sum_{i=1}^n \frac{-4\lambda y_i^{2\phi}}{4\lambda^4} = \frac{-n}{\lambda} + \frac{\sum_{i=1}^n y_i^{2\phi}}{\lambda^3} := 0$$

or

$$\lambda^2 = \frac{\sum_{i=1}^n y_i^{2\phi}}{n}, or \hat{\lambda} = \sqrt{\frac{\sum_{i=1}^n y_i^{2\phi}}{n}}.$$

Notice that

$$\frac{d^2\log L}{d\lambda^2}|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \frac{\sum_{i=1}^n 3y_i^{2\phi}}{\hat{\lambda}^4} = \frac{n}{\hat{\lambda}^2}[1-3] < 0.$$

Hence $\hat{\lambda}$ is the MLE of λ if ϕ is known.

CHAPTER 3

SIMULATIONS COVERAGE OF CONFIDENCE INTERVALS

3.1 HALF-NORMAL DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for μ and σ^2 with nominal 90% and 95% confidence levels respectively and sample sizes ranging from 5 to 500. Two types of confidence intervals for μ are used; the Pewsey interval is given by

$$(Y_{(1)} + \hat{\sigma}\log(\frac{\alpha}{2})\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right), Y_{(1)} + \hat{\sigma}\log(1-\frac{\alpha}{2})\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right)) \tag{3.1}$$

and our new modified confidence interval

$$(Y_{(1)} + \hat{\sigma}\log(\alpha)\Phi^{-1}\left(\frac{1}{2}+\frac{1}{2n}\right)\left(1+\frac{13}{n^2}\right), Y_{(1)}). \tag{3.2}$$

The confidence interval used for σ^2 when μ is known is given by

$$(\frac{T_n}{\chi^2_{n,1-\frac{\alpha}{2}}}, \frac{T_n}{\chi^2_{n,\frac{\alpha}{2}}}) \tag{3.3}$$

where $T_n = \sum_{i=1}^{n}(Y_i - \mu)^2$.

and when μ is unknown is given by

$$(\frac{D_n}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{D_n}{\chi^2_{n-1,\frac{\alpha}{2}}}) \tag{3.4}$$

where $D_n = \sum_{i=1}^{n}(Y_i - Y_{(1)})^2$ Each quoted coverage percentage was obtained from 5000 pseudo-random samples of size n from the (standard) half-normal distribution. The standard error of any entry is thus, at most, 0.004. Comparing the coverage of Pewsey interval and the modified Pewsey interval for μ we notice that the modified Pewsey interval has higher coverage for small sample sizes (n $<$ 10) and similar coverage to that of Pewsey interval for other sample sizes, also, we note that it has a shorter length except when n =5 where the coverage is better.

Table 3.1. Actual Coverage Levels for Nominal 90% Confidence Interval for σ^2 when μ is unknown for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.8914	9.83
10	9	0.8856	6.30
25	24	0.8964	5.15
50	49	0.8948	4.89
100	99	0.8980	4.76
500	499	0.8984	4.68
∞	∞	.9	4.65

The confidence interval formula used is (3.4).

Table 3.2. Actual Coverage Levels for Nominal 90% Confidence Interval for σ^2 when μ is known for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.9000	8.73
10	9	0.8998	6.28
25	24	0.9010	5.22
50	49	0.8992	4.94
100	99	0.8990	4.79
500	499	0.8978	4.68
∞	∞	.9	4.65

The confidence interval formula used is (3.3).

Table 3.3. Actual Coverage Levels for Nominal 90% Confidence Interval for μ for sample sizes ranging from 5 to 500 - Modified and Pewsey intervals

		modified		Pewsey	
n	d = (n-1)	Coverage	Slen	Coverage	Slen
5	4	0.9158	3.40	0.8410	2.86
10	9	0.8948	2.88	0.8694	3.26
25	24	0.8930	2.80	0.9000	3.51
50	49	0.8920	2.83	0.8898	3.60
100	99	0.9000	2.85	0.9026	3.64
500	499	0.9020	2.88	0.9052	3.68
∞	∞	.9	2.89	.9	3.69

The confidence interval formulas used are (3.1) and (3.2).

Table 3.4. Actual Coverage Levels for Nominal 95% Confidence Interval for σ^2 when μ is unknown for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.9438	15.46
10	9	0.9470	8.16
25	24	0.9450	6.37
50	49	0.9488	5.95
100	99	0.9478	5.71
500	499	0.9448	5.58
∞	∞	.95	5.54

The confidence interval formula used is (3.4).

Table 3.5. Actual Coverage Levels for Nominal 95% Confidence Interval for σ^2 when μ is known for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.9508	12.52
10	9	0.9528	9.92
25	24	0.9476	6.89
50	49	0.9476	6.19
100	99	0.9500	5.82
500	499	0.9458	5.60
∞	∞	.95	5.54

The confidence interval formula used is (3.3).

Table 3.6. Actual Coverage Levels for Nominal 95% Confidence Interval for μ for sample sizes ranging from 5 to 500 - Modified and Pewsey intervals

		modified		Pewsey	
n	d = (n-1)	Coverage	Slen	Coverage	Slen
5	4	0.9474	4.44	0.9004	3.57
10	9	0.9404	3.75	0.9320	4.03
25	24	0.9438	3.64	0.9438	4.36
50	49	0.9478	3.69	0.9480	4.48
100	99	0.9482	3.71	0.9496	4.53
500	499	0.9456	3.74	0.9474	4.58
∞	∞	.95	3.76	.95	4.59

The confidence interval formulas used are (3.1) and (3.2).

3.2 HPARETO DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for θ and λ^2 with nominal 95% confidence level and sample sizes ranging from 5 to 500. Each quoted coverage percentage was obtained from 5000 pseudo-random samples of size n from the (standard) half-normal distribution. The standard error of any entry is thus, at most, 0.004.

Table 3.7. Actual Coverage Levels for Nominal 95% Confidence Interval for θ for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.8920	3.54
10	9	0.9220	4.09
25	24	0.944	4.38
50	49	0.9496	4.47
100	99	0.9504	4.53
500	499	0.9590	4.58
∞	∞	.95	4.59

The confidence interval formula used is (2.15).

Table 3.8. Actual Coverage Levels for Nominal 95% Confidence Interval for λ^2 when θ is unknown for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.9414	14.76
10	9	0.9460	8.42
25	24	0.9468	6.42
50	49	0.9460	5.90
100	99	0.9498	5.73
500	499	0.9522	5.67
∞	∞	.95	5.54

The confidence interval formula used is (2.13).

Table 3.9. Actual Coverage Levels for Nominal 95% Confidence Interval for λ^2 when θ is known for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.9458	21.95
10	9	0.9458	10.20
25	24	0.9500	6.940
50	49	0.9494	6.13
100	99	0.9508	5.85
500	499	0.9532	5.59
∞	∞	.95	5.54

The confidence interval formula used is (2.14).

3.3 HPOWER DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for λ^2 with nominal 95% confidence level and sample sizes ranging from 5 to 500. Each quoted coverage percentage was obtained from 5000 pseudo-random samples of size n from the (standard) half-normal distribution. The standard error of any entry is thus, at most, 0.004. The confidence interval formula used is (2.18).

Table 3.10. Actual Coverage Levels for Nominal 95% Confidence Interval for λ^2 for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.9510	21.90
10	9	0.9500	10.07
25	24	0.9556	6.90
50	49	0.9538	6.17
100	99	0.9482	5.85
500	499	0.9475	5.61
∞	∞	.95	5.54

3.4 HTEV DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for λ^2 with nominal 95% confidence level and sample sizes ranging from 5 to 500. Each quoted coverage percentage was obtained from 5000 pseudo-random samples of size n from the (standard) half-normal distribution. The standard error of any entry is thus, at most, 0.004. The confidence interval formula used is (2.26).

Table 3.11. Actual Coverage Levels for Nominal 95% Confidence Interval for λ^2 for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.9426	22.30
10	9	0.9438	10.10
25	24	0.9496	6.91
50	49	0.9490	6.17
100	99	0.9510	5.84
500	499	0.9494	5.60
∞	∞	.95	5.54

3.5 TWO PARAMETER EXPONENTIAL DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for θ and λ^2 with nominal 95% confidence level and sample sizes ranging from 5 to 500. Each quoted coverage percentage was obtained from 5000 pseudo-random samples of size n from the exponential distribution. The standard error of any entry is thus, at most, 0.004. The confidence interval formula used is (2.2).

Table 3.12. Actual Coverage Levels for Nominal 95% Confidence Interval for λ when θ is unknown for sample sizes ranging from 5 to 500

n	d = 2(n-1)	Coverage	Slen
5	4	0.9520	7.11
10	9	0.9452	5.08
25	24	0.9518	4.34
50	49	0.9470	4.11
100	99	0.9484	4.02
500	499	0.9496	3.93
∞	∞	.95	3.92

Table 3.13. Actual Coverage Levels for Nominal 95% Confidence Interval for λ when θ is known for sample sizes ranging from 5 to 500

n	d = 2n	Coverage	Slen
5	4	0.9536	5.74
10	9	0.9480	4.72
25	24	0.9518	4.24
50	49	0.9464	4.07
100	99	0.9482	4.00
500	499	0.9496	3.93
∞	∞	.95	3.92

The confidence interval formula used is (2.3).

Table 3.14. Actual Coverage Levels for Nominal 95% Confidence Interval for θ for sample sizes ranging from 5 to 500

n	d = (n-1)	Coverage	Slen
5	4	0.9494	4.41
10	9	0.9460	3.53
25	24	0.9466	3.20
50	49	0.9466	3.09
100	99	0.9496	3.04
500	499	0.9514	3.00
∞	∞	.95	3.00

The confidence interval formula used is (2.1).

3.6 THE PARETO DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for σ and λ with nominal 95% confidence level and sample sizes ranging from 5 to 500. Each quoted coverage percentage was obtained from 5000 pseudo-random samples of size n from the pareto distribution. The standard error of any entry is thus, at most, 0.004.

The confidence interval formula used is (2.12).

Table 3.15. Actual Coverage Levels for Nominal 95% Confidence Interval for λ when σ is unknown for sample sizes ranging from 5 to 500

n	d = 2(n-1)	Coverage	Slen
5	4	0.9490	7.28
10	9	0.9464	5.12
25	24	0.9522	4.34
50	49	0.9474	4.11
100	99	0.9486	4.01
500	499	0.9466	3.94
∞	∞	.95	3.92

Table 3.16. Actual Coverage Levels for Nominal 95% Confidence Interval for σ for sample sizes ranging from 5 to 500

n	d = n-1	Coverage	Slen
5	4	0.9496	3.48
10	9	0.9432	3.28
25	24	0.9510	3.12
50	49	0.9482	3.06
100	99	0.9546	3.02
500	499	0.9518	3.00
∞	∞	.95	3.00

The confidence interval formula used is (2.11).

3.7 THE POWER DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for λ with nominal 95% confidence level and sample sizes ranging from 5 to 500. Each quoted coverage percentage was obtained from 5000 pseudo-random samples of size n from the Power distribution. The standard error of any entry is thus, at most, 0.004.

The confidence interval formula used is (2.17).

Table 3.17. Actual Coverage Levels for Nominal 95% Confidence Interval for λ for sample sizes ranging from 5 to 500

n	d = 2n	Coverage	Slen
5	4	0.9464	5.80
10	9	0.9540	4.75
25	24	0.9506	4.23
50	49	0.9500	4.06
100	99	0.9458	4.00
500	499	0.9442	3.93
∞	∞	.95	3.92

3.8 THE TRUNCATED EXTREME VALUE DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for λ with nominal 95% confidence level and sample sizes ranging from 5 to 500. Each quoted coverage percentage was obtained from 5000 pseudo-random samples of size n from the truncated extreme value distribution. The standard error of any entry is thus, at most, 0.004. The confidence interval formula used is (2.25).

Table 3.18. Actual Coverage Levels for Nominal 95% Confidence Interval for λ for sample sizes ranging from 5 to 500

n	d = 2n	Coverage	Slen
5	4	0.9454	5.83
10	9	0.9474	4.73
25	24	0.9526	4.22
50	49	0.9574	4.08
100	99	0.9552	4.00
500	499	0.9496	3.93
∞	∞	.95	3.92

3.9 THE WEIBULL DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for ϕ and λ with nominal 95% confidence level and sample sizes ranging from 25 to 500. Each quoted coverage percentage was obtained from 100 pseudo-random samples of size n from the Weibull distribution. The standard error of any entry is thus, at most, 0.043. The confidence interval formulas used are (2.27) and (2.28).

Table 3.19. Actual Coverage Levels for Nominal 95% Confidence Interval for ϕ and λ for sample sizes ranging from 25 to 500

	$\phi = 1$		$\lambda = 1$	
n	Coverage	Slen	Coverage	Slen
25	.95	3.29	.94	4.35
50	.91	3.12	.94	4.23
100	.94	3.05	.92	4.18
500	.96	3.07	.94	4.15
∞	.95	3.06	.95	4.13

Table 3.20. Actual Coverage Levels for Nominal 95% Confidence Interval for ϕ and λ for sample sizes ranging from 25 to 500

	$\phi = 1$		$\lambda = 5$	
n	Coverage	Slen	Coverage	Slen
25	.95	3.23	.93	48.56
50	.92	3.16	.94	43.44
100	.93	3.07	.90	38.99
500	.94	3.06	.95	37.65
∞	.95	3.06	.95	36.73

The confidence interval formulas used are (2.27) and (2.28).

Table 3.21. Actual Coverage Levels for Nominal 95% Confidence Interval for ϕ and λ for sample sizes ranging from 25 to 500

	$\phi = 1$		$\lambda = 10$	
n	Coverage	Slen	Coverage	Slen
25	.95	3.22	.97	133.83
50	.97	3.16	.97	115.44
100	.95	3.12	.96	107.08
500	.94	3.07	.96	94.45
∞	.95	3.06	.95	92.07

The confidence interval formulas used are (2.27) and (2.28).

Table 3.22. Actual Coverage Levels for Nominal 95% Confidence Interval for ϕ and λ for sample sizes ranging from 25 to 500

	$\phi = 1$		$\lambda = 20$	
n	Coverage	Slen	Coverage	Slen
25	.94	3.05	.91	332.56
50	.96	3.12	.92	273.23
100	.95	3.14	.95	270.37
500	.97	3.06	.96	231.96
∞	.95	3.06	.95	223.20

The confidence interval formulas used are (2.27) and (2.28).

Table 3.23. Actual Coverage Levels for Nominal 95% Confidence Interval for ϕ and λ for sample sizes ranging from 25 to 500

	$\phi = 20$		$\lambda = 1$	
n	Coverage	Slen	Coverage	Slen
25	.98	63.63	.95	4.14
50	.90	62.93	.91	4.23
100	.94	62.20	.93	4.22
500	.96	61.51	.98	4.16
∞	.95	61.13	.95	4.13

The confidence interval formulas used are (2.27) and (2.28).

Table 3.24. Actual Coverage Levels for Nominal 95% Confidence Interval for ϕ and λ for sample sizes ranging from 25 to 500

	$\phi = 20$		$\lambda = 5$	
n	Coverage	Slen	Coverage	Slen
25	.99	65.17	.97	47.18
50	.97	62.56	.94	43.34
100	.93	62.69	.94	41.11
500	.98	61.53	1.00	37.80
∞	.95	61.13	.95	36.73

The confidence interval formulas used are (2.27) and (2.28).

Table 3.25. Actual Coverage Levels for Nominal 95% Confidence Interval for ϕ and λ for sample sizes ranging from 25 to 500

	$\phi = 20$		$\lambda = 10$	
n	Coverage	Slen	Coverage	Slen
25	.94	61.78	.92	135.83
50	.94	61.75	.91	106.16
100	.97	61.34	.97	97.01
500	.96	61.15	.95	94.72
∞	.95	61.13	.95	92.07

The confidence interval formulas used are (2.27) and (2.28).

Table 3.26. Actual Coverage Levels for Nominal 95% Confidence Interval for ϕ and λ for sample sizes ranging from 25 to 500

	$\phi = 20$		$\lambda = 20$	
n	Coverage	Slen	Coverage	Slen
50	.92	61.61	.91	261.31
100	.94	61.56	.95	248.74
500	.94	61.51	.95	237.89
∞	.95	61.13	.95	223.20

The confidence interval formulas used are (2.27) and (2.28).

3.10 THE RAYLEIGH DISTRIBUTION

The following tables show the results obtained from a simulation study designed to establish the actual coverage of confidence intervals for μ and σ with nominal 95% confidence level and sample sizes ranging from 25 to 500. Each quoted coverage percentage was obtained from 100 pseudo-random samples of size n from the Rayleigh distribution. The standard error of any entry is thus, at most, 0.043. The confidence interval formulas used are (2.22) and (2.23).

Table 3.27. Actual Coverage Levels for Nominal 95% Confidence Interval for μ and σ for sample sizes ranging from 25 to 500

	$\mu = 1$		$\sigma = 1$	
n	Coverage	Slen	Coverage	Slen
25	.91	2.84	.94	2.81
50	.98	2.35	.96	2.47
100	.93	2.22	.93	3.04
500	.92	2.32	.94	2.24

Table 3.28. Actual Coverage Levels for Nominal 95% Confidence Interval for μ and σ for sample sizes ranging from 25 to 500

	$\mu = 2$		$\sigma = 5$	
n	Coverage	Slen	Coverage	Slen
25	.92	13.58	.92	11.82
50	.93	12.06	.94	12.24
100	.92	12.06	.96	12.42
500	.94	9.99	.94	12.97

Table 3.29. Actual Coverage Levels for Nominal 95% Confidence Interval for μ and σ for sample sizes ranging from 25 to 500

	$\mu = 2$		$\sigma = 10$	
n	Coverage	Slen	Coverage	Slen
25	.91	29.29	.94	26.71
50	.93	22.78	.94	24.74
100	.92	22.31	.96	23.16
500	.93	23.31	.95	23.63

Table 3.30. Actual Coverage Levels for Nominal 95% Confidence Interval for μ and σ for sample sizes ranging from 25 to 500

	$\mu = 5$		$\sigma = 2$	
n	Coverage	Slen	Coverage	Slen
25	.86	4.97	.99	24.82
50	.91	4.65	.95	4.46
100	.92	4.54	.92	5.25
500	.92	4.22	.96	4.65

Table 3.31. Actual Coverage Levels for Nominal 95% Confidence Interval for μ and σ for sample sizes ranging from 25 to 500

	$\mu = 10$		$\sigma = 2$	
n	Coverage	Slen	Coverage	Slen
25	.92	5.43	.96	5.22
50	.92	4.30	.93	4.91
100	.95	4.89	.95	4.79
500	.92	4.21	.97	4.43

Table 3.32. Actual Coverage Levels for Nominal 95% Confidence Interval for μ and σ for sample sizes ranging from 25 to 500

	$\mu = 20$		$\sigma = 20$	
n	Coverage	Slen	Coverage	Slen
25	.92	57.34	.93	49.13
50	.87	44.79	.94	50.50
100	.93	42.22	.96	46.10
500	.93	37.55	.96	40.07

The following tables show the results obtained from a simulation study designed to find the sample means and standard deviations of the MLEs for μ and σ using sample sizes ranging from 25 to 500. Each quoted mean and standard deviation was obtained from 100 pseudo-random samples of size n from the Rayleigh distribution. The MLEs were found by iteration using Newton's method, and the number of iterations used were 100.

Table 3.33. Sample means and standard deviations of the MLEs for μ and σ using sample sizes ranging from 25 to 500

	$\mu = 1$		$\sigma = 1$	
n	$\hat{\mu}$	$SD(\hat{\mu})$	$\hat{\sigma}$	$SD(\hat{\sigma})$
25	1.08	0.14	0.95	0.14
50	1.03	0.09	0.98	0.09
100	1.02	0.07	1.02	0.07
500	1.01	0.03	1.02	0.03

Table 3.34. Sample means and standard deviations of the MLEs for μ and σ using sample sizes ranging from 25 to 500

	$\mu = 2$		$\sigma = 5$	
n	$\hat{\mu}$	$SD(\hat{\mu})$	$\hat{\sigma}$	$SD(\hat{\sigma})$
25	2.34	0.69	4.69	0.60
50	2.23	0.43	4.76	0.44
100	2.14	0.31	4.90	0.32
500	2.02	0.11	5.00	0.15

Table 3.35. Sample means and standard deviations of the MLEs for μ and σ using sample sizes ranging from 25 to 500

	$\mu = 2$		$\sigma = 10$	
n	$\hat{\mu}$	$SD(\hat{\mu})$	$\hat{\sigma}$	$SD(\hat{\sigma})$
25	3.01	1.49	9.27	1.36
50	2.39	0.82	9.72	0.89
100	2.19	0.57	9.78	0.59
500	2.05	0.27	10.00	0.27

Table 3.36. Sample means and standard deviations of the MLEs for μ and σ using sample sizes ranging from 25 to 500

	$\mu = 5$		$\sigma = 2$	
n	$\hat{\mu}$	$SD(\hat{\mu})$	$\hat{\sigma}$	$SD(\hat{\sigma})$
25	5.19	0.25	1.83	0.24
50	5.03	0.17	1.96	0.16
100	5.04	0.12	1.97	0.13
500	5.02	0.05	1.99	0.05

Table 3.37. Sample means and standard deviations of the MLEs for μ and σ using sample sizes ranging from 25 to 500

	$\mu = 10$		$\sigma = 2$	
n	$\hat{\mu}$	$SD(\hat{\mu})$	$\hat{\sigma}$	$SD(\hat{\sigma})$
25	10.10	0.27	1.94	0.27
50	10.07	0.16	1.93	0.18
100	10.03	0.13	1.99	0.12
500	10.02	0.05	1.99	0.05

Table 3.38. Sample means and standard deviations of the MLEs for μ and σ using sample sizes ranging from 25 to 500

	$\mu = 20$		$\sigma = 20$	
n	$\hat{\mu}$	$SD(\hat{\mu})$	$\hat{\sigma}$	$SD(\hat{\sigma})$
25	21.57	2.93	19.13	2.51
50	20.93	1.82	19.27	1.82
100	20.44	1.07	19.68	1.18
500	20.13	0.43	19.94	0.46

REFERENCES

[1] Arnold, B. C. (1983), *Pareto Distributions*, Co-operative Publishing House, Burtonsville, Maryland.

[2] Bain, L. J. (1978), *Statistical Analysis of Reliability and Life-Testing Models*, Marcel Dekker, NY.

[3] Bury, K. (1999), *Statistical Distributions in Engineering*, Cambridge University Press, Cambridge, UK.

[4] Casella, G., and Berger, R. L. (2002), *Statistical Inference*, Duxbury, Pacific Grove, CA.

[5] Cohen, A. C. (1988), *Parameter Estimation in Reliability and Life Span Models*, Marcel Dekker, NY.

[6] Evans, M. (2000), *Statistical Distributions*, John Wiley and Sons, NY.

[7] Johnson, N.L., and Kotz, S. (1970), *Continuous Univariate Distributions*, vol. 1, Houghton Mifflin Company, Boston, MA.

[8] Krishnamorthy, K. (2006), *Handbook of Statistical Distributions with Applications*, Taylor & Francis Group, Broken Sound Parkway, NW.

[9] Lehmann, E. L. (1983), *Theory of Point Estimation*, John Wiley and Sons, NY.

[10] Mann, N.R., Schafer, R. E., and Singpurwalla, N.D. (1974), *Methods for Statistical Analysis of Reliability and Life Data*, John Wiley and Sons, NY.

[11] Meeker, W. Q., Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, John Wiley and Sons, NY.

[12] Olive, D.J. (2007), *A Course in Statistical Theory*, available from (http://www.math.siu.edu/olive/infbook.htm).

[13] Olive, D.J. (2007), "A Simple Limit Theorem for Exponential Families," available from (http://www.math.siu.edu/olive).

[14] Olive, D.J. (2007), "Robust Estimators for Transformed Location Scale Fami-

lies," available from (http://www.math.siu.edu/olive).

[15] Pewsey, A. (2002), "Large-Sample Inference for the General Half-Normal Distribution," *Communications in Statistics Theory and Methods*, 31, 1045-1054.

APPENDIX

R CODE FOR THE SIMULATIONS, the Statistical package is available from (http://www.r-project.org/).

```
 #simulates exp 100(1-alpha)% CI for lambda and CI for theta,
 expsim<-function(n = 10, nruns = 5000, theta = 0, lambda = 1,
alpha = 0.05, p = 1)
{ scov <- 0
ccov <- 0
mcov <- 0
slow <- 1:nruns
sup <- slow
mlow <- slow
mup <- slow
clow <- slow
cup <- slow
ucut <- alpha/2
lcut <- 1 - ucut
d <- 2 * (n - p)
d2 <- 2 * n
lval <- log(alpha/2)/n
uval <- log(1 - alpha/2)/n
mlval <- alpha^(-1/(n - 1)) - 1
for(i in 1:nruns) {
y <- theta + lambda * rexp(n)
miny <- min(y)
dn <- sum((y - miny))
```

```
        #get CI for lambda
lamhat <- dn/n
num <- 2 * dn
slow[i] <- num/qchisq(lcut, df = d)
sup[i] <- num/qchisq(ucut, df = d)
if(slow[i] < lambda && sup[i] > lambda) scov <- scov + 1
#get CI for lambda when theta is known
tn <- sum((y - theta))
num <- 2 * tn
clow[i] <- num/qchisq(lcut, df = d2)
cup[i] <- num/qchisq(ucut, df = d2)
if(clow[i] < lambda && cup[i] > lambda)
ccov <- ccov + 1
#get  CI for theta
mlow[i] <- miny - lamhat * mlval
mup[i] <- miny
if(mlow[i] < theta) mcov <- mcov + 1
}
scov <- scov/nruns
slen <- sqrt(n) * mean(sup - slow)
mcov <- mcov/nruns
mlen <- n * mean(mup - mlow)
ccov <- ccov/nruns
clen <- sqrt(n) * mean(cup - clow)
list(d = d, scov = scov, slen = slen,
ccov = ccov, clen = clen, mcov = mcov, mlen = mlen)
}
```

```
#simulates Pewsey HN 100(1-alpha)% CI for sigma^2 and one for mu,
#The CI for mu should work better than the Pewsey interval.
hnsim<-
function(n = 10, nruns = 5000, mu = 0, sigma = 1, alpha = 0.05, p = 1)
{
scov <- 0
lcov <- 0
lcov2 <- 0
ccov <- 0
slow <- 1:nruns
sup <- slow
llow <- slow
lup <- slow
llow2 <- slow
lup2 <- slow
clow <- slow
cup <- sup
ucut <- alpha/2
lcut <- 1 - ucut
sigsq <- sigma^2
d <- n - p
phiinv <- qnorm((0.5 + 1/(2 * n)))
lval <- log(alpha) * phiinv * (1 + 13/n^2)
lval2 <- log(alpha/2) * phiinv
lval3 <- log(1-alpha/2) * phiinv
for(i in 1:nruns) {
y <- mu + sigma * abs(rnorm(n))
```

```
miny <- min(y)
dn <- sum((y - miny)^2)
#get CI for sigma^2 when mu is unknown
slow[i] <- dn/qchisq(lcut, df = d)
sup[i] <- dn/qchisq(ucut, df = d)
if(slow[i] < sigsq && sup[i] > sigsq)
scov <- scov + 1
#get CI for sigma^2 if mu is known
tn <- sum((y - mu)^2)
clow[i] <- tn/qchisq(lcut, df = n)
cup[i] <- tn/qchisq(ucut, df = n)
if(clow[i] < sigsq && cup[i] > sigsq)
ccov <- ccov + 1
#get CI for mu (modified Pewsey type interval)
shat <- sqrt(dn/n)
llow[i] <- miny + shat * lval
lup[i] <- miny
if(llow[i] < mu)
lcov <- lcov + 1
#get CI for mu (Pewsey type interval)
shat <- sqrt(dn/n)
llow2[i] <- miny + shat * lval2
lup2[i] <- miny + shat * lval3
if(llow2[i] < mu && lup2[i] > mu)
lcov2 <- lcov2 + 1
}
scov <- scov/nruns
```

```
slen <- sqrt(n) * mean(sup - slow)
lcov <- lcov/nruns
llen <- n * mean(lup - llow)
lcov2 <- lcov2/nruns
llen2 <- n * mean(lup2 - llow2)
ccov <- ccov/nruns
clen <- sqrt(n) * mean(cup - clow)
list(d = d, scov = scov, slen = slen, ccov =
ccov, clen = clen,lcov = lcov, llen = llen, lcov2 = lcov2, llen2 = llen2,)
}
```

```
#simulates HPareto 100(1-alpha)% CI for sigma^2 and one for mu,
 "hparsim1"<-
function(n = 10, nruns = 5000, mu = 0, sigma = 1, alpha = 0.05, p = 1)
{
scov <- 0
lcov <- 0
ccov <- 0
slow <- 1:nruns
sup <- slow
llow <- slow
lup <- slow
clow <- slow
cup <- sup
ucut <- alpha/2
lcut <- 1 - ucut
sigsq <- sigma^2
d <- n - p
phiinv <- qnorm((0.5 + 1/(2 * n)))
lval <- log(alpha/2) * phiinv
uval <- log(1 - alpha/2) * phiinv
for(i in 1:nruns) {
w <- mu + sigma * abs(rnorm(n))
                y<-exp(w)
                minw<- min(w)
miny <- min(y)
dn <- sum((log(y) - log(miny))^2)
                dn1 <- sum((log(y) - mu)^2)
```

```
#get CI for sigma^2
slow[i] <- dn/qchisq(lcut, df = d)
sup[i] <- dn/qchisq(ucut, df = d)
if(slow[i] < sigsq && sup[i] > sigsq) scov <- scov + 1
#get CI for mu
shat <- sqrt(dn/n)
llow[i] <- minw + shat * lval
lup[i] <- minw + shat * uval
if(llow[i] < mu && lup[i] > mu) lcov <- lcov + 1
#get CI for sigma^2 if mu is known
clow[i] <- dn1/qchisq(lcut, df = d)
cup[i] <- dn1/qchisq(ucut, df = d)
if(clow[i] < sigsq && cup[i] > sigsq)
ccov <- ccov + 1
}
scov <- scov/nruns
slen <- sqrt(n) * mean(sup - slow)
lcov <- lcov/nruns
llen <- n * mean(lup - llow)
ccov <- ccov/nruns
clen <- sqrt(n) * mean(cup - clow)
list(d = d, scov = scov, slen = slen, lcov = lcov, llen = llen,
ccov = ccov, clen = clen,  )
}
```

```
#Simulates  HPower 100(1-alpha)% CI for sigma^2,
hpowsim<-function(n, nruns=5000, mu=0, sigma=1, alpha=0.05, d=n-1)
{           cov <- 0
            low <- 1:nruns
            up <- low
            ucut <- alpha/2
            lcut <- 1-ucut
            sigsq <-sigma^2
            for(i in 1:nruns){
                        w<-mu + sigma * abs(rnorm(n))
                        y<- exp(-w)
                        miny<-min(y)
                        wn<- sum((log(y))^2)
                        low[i] <- wn/qchisq(lcut,df=d)
                        up[i] <- wn/qchisq(ucut, df= d)
                        if(low[i] < sigsq && up[i] > sigsq)
                                       cov <- cov + 1
          }
          cov <- cov / nruns
          slen <- sqrt(n) * mean(up - low)
          list(d = d, cov = cov, slen = slen)
}
```

```
#Simulates HTEV 100(1-alpha)% CI for sigma^2,
 htevsim<-function(n, nruns=5000, mu=0, sigma=1, alpha=0.05, d=n-1)
 {
            cov <- 0
            low <- 1:nruns
            up <- low
            ucut <- alpha/2
            lcut <- 1-ucut
            sigsq <-sigma^2
            for(i in 1:nruns){
                        w<-mu + sigma * abs(rnorm(n))
                        y<-log(w+1)
                        wn<- sum((exp(y)-1)^2)
                        low[i] <- wn/qchisq(lcut,df=d)
                        up[i] <- wn/qchisq(ucut, df= d)
                        if(low[i] < sigsq && up[i] > sigsq)
                                       cov <- cov + 1
          }
          cov <- cov / nruns
          slen <- sqrt(n) * mean(up - low)
          list(d = d, cov = cov, slen = slen)
 }
```

```
#simulates Pareto 100(1-alpha)% CI for lambda and CI for theta,
parsim<-function(n = 10, nruns = 5000, theta = 0, lambda = 1, alpha = 0.05, p
= 1)
{ scov <- 0
ccov <- 0
mcov <- 0
slow <- 1:nruns
sup <- slow
mlow <- slow
mup <- slow
clow <- slow
cup <- slow
ucut <- alpha/2
lcut <- 1 - ucut
d <- 2 * (n - p)
d2 <- 2 * n
lval <- log(alpha/2)/n
uval <- log(1 - alpha/2)/n
mlval <- alpha^(-1/(n - 1)) - 1
sigma <- exp(theta)
for(i in 1:nruns) {
w <- theta + lambda * rexp(n)
y <- exp(w)
minw <- min(w)
dn <- sum((w - minw))
        #get CI for lambda
lamhat <- dn/n
```

```
num <- 2 * dn
slow[i] <- num/qchisq(lcut, df = d)
sup[i] <- num/qchisq(ucut, df = d)
if(slow[i] < lambda && sup[i] > lambda) scov <- scov + 1
#get CI for theta
mlow[i] <- exp(minw - lamhat * mlval)
mup[i] <- exp(minw)
if(mlow[i] < sigma) mcov <- mcov + 1
}
scov <- scov/nruns
slen <- sqrt(n) * mean(sup - slow)
mcov <- mcov/nruns
mlen <- n * mean(mup - mlow)
ccov <- ccov/nruns
clen <- sqrt(n) * mean(cup - clow)
list(d = d, scov = scov, slen = slen, mcov = mcov, mlen = mlen)
}
```

```
#Simulates POW 100(1-alpha)% CI for lambda^2,
powsim<-function(n, nruns=5000, mu=0, lam=1, alpha=0.05, d=2*n)
{       theta = 0
            cov <- 0
            low <- 1:nruns
            up <- low
            ucut <- alpha/2
            lcut <- 1-ucut
            lamsq <-lam^2
            for(i in 1:nruns){
                        w<-theta + lam * rexp(n)
                        y<-exp(-w)
                        wn<- 2 * sum(log(1/y))
                        low[i] <- wn/qchisq(lcut,df=d)
                        up[i] <- wn/qchisq(ucut, df= d)
                        if(low[i] < lamsq && up[i] > lamsq)
                                       cov <- cov + 1
          }
          cov <- cov / nruns
          slen <- sqrt(n) * mean(up - low)
          list(d = d, cov = cov, slen = slen)
}
```

```
#simulates MLEs and CIs for mu and sigma in the Rayleigh distribution
 raysim <- function(n = 100, mu = 1, sigma = 1, runs = 100, iter = 100)
  {  countm <- 0
counts <- 0
count2s <- 0
munew <- 1:runs
     sigmanew <- munew
sigmanew2 <- munew
muo <- 1:runs
sigmao <- 1:runs
meanw <- 1:runs
meany <- 1:runs
mnew <- 0
snew <- 0
muold <- 0
sigmaold <- 0
muvold <- 0
sigmavold <- 0
    vec <- 1:6
mcov <- 0
scov <- 0
mlow <- 0
mup <- 0
slow <- 0
sup <- 0
    for(i in 1:runs) {
        w <- 2 * sigma^2 * rexp(n)
```

```
        y <- sqrt(w) + mu
sigmaold <- sqrt(var(y)/0.429204)
muold <- mean(y)-1.25331 * sigmaold
muold <- min(muold,2*min(y)-muold)
mvvold <- muold
sigmaold <- sqrt(sum((y-muold)^2)/(2 * n))
svvold <- sigmaold
for(j in 1:iter) {
D <- -2*n/sigmaold^2*sum((y-muold)^-2)+3/sigmaold^4*sum((y-muold)^-2)
*sum((y-muold)^2)- 2*n^2/sigmaold^4+3*n/sigmaold^6*sum((y-muold)^2)
-4/sigmaold^6*(sum(y-muold))^2
a1 <- -2*n/sigmaold^2*sum((y-muold)^-1)+2*n/sigmaold^4*sum(y-muold)
+3*sum((y-muold)^2)* sum((y-muold)^-1)/sigmaold^4-3*sum((y-muold)^2)
*sum(y-muold)/sigmaold^6- 4*n*sum(y-muold)/sigmaold^4+2*sum(y-muold)
*sum((y-muold)^2)/sigmaold^6
a2 <- - 2/sigmaold^3*sum(y-muold)*sum((y-muold)^-1)+2/sigmaold^5
*(sum(y-muold))^2+ 2*n^2/sigmaold^3-n*sum((y-muold)^2)/sigmaold^5
+2*n/sigmaold*sum((y-muold)^-2)- 1/sigmaold^3*sum((y-muold)^-2)
*sum((y-muold)^2)
snew <-  sigmaold - (a2/D)
mnew <-  muold - (a1/D)
mnew <- min(mnew, 2*min(y)-mnew)
if(mnew < min(y)/100){countm <- countm +1}
if(snew < 0){counts <- counts +1}
if(snew > 10*sigma){count2s <- count2s +1}
if(mnew < min(y)/100){mnew <- min(y)-.01}
if(snew < 0){snew <- sigmaold}
```

```
if(snew > 10*sigma){snew <- sigmaold}
muvold <- muold
muold <- mnew
sigmavold <- sigmaold
sigmaold <- snew
}
muo[i] <- muvold
sigmao[i] <- sigmavold
munew[i] <- mnew
sigmanew[i] <- snew # Iteration formula
sigmanew2[i] <- sqrt(sum((y-munew[i])^2)/(2 * n))#Exact formula
Sig11 <- (1/D)*[2*n/sigmanew[i]^2 -3/sigmanew[i]^4 * sum((y-munew[i])^2)]
Sig22 <- (-1/D)*[sum((y-munew[i])^-2)+ n/sigmanew[i]^2]
SDmuI <- sqrt(Sig11)
SDsigI <- sqrt(Sig22)
#get CI for mu
mlow[i] <- munew[i] - 1.96 * sqrt(Sig11)
mup [i] <- munew[i] + 1.96 * sqrt(Sig11)
if(mlow[i] < mu && mup[i] > mu) {mcov <- mcov + 1}
#get CI for sigma
slow[i] <- sigmanew[i] - 1.96 * sqrt(Sig22)
sup [i] <- sigmanew[i] + 1.96 * sqrt(Sig22)
if(slow[i] < sigma && sup[i] > sigma) {scov <- scov + 1}
  }
mcov <- mcov/runs
scov <- scov/runs
slenm <- sqrt(n) * mean(mup - mlow)
```

```
slens <- sqrt(n) * mean(sup - slow)
 vec[1] <- mean(munew)
 vec[2] <- sqrt(var(munew))
 vec[3] <- mean(sigmanew)
 vec[4] <-  sqrt(var(sigmanew))
 vec[5] <- mean(sigmanew2)
 vec[6] <- sqrt(var(sigmanew2))
mconv <- max(abs(munew - muo))
sconv <- max(abs(sigmanew - sigmao))
list(y=y,mvvold=mvvold,svvold=svvold,mu = mu, mumle = munew,sigma = sigma,
sigmamle = sigmanew,sigmaEX = sigmanew2, meanmuN = vec[1],SDmuN = vec[2],
SDmuI = SDmuI, meansigmaN = vec[3],SDsigN = vec[4], SDsigI = SDsigI,
meansigmaEx = vec[5],SDsigEx = vec[6],
muconv = mconv,CIlengthmu=slenm,CILengthsig=slens,sigmaconv = sconv,countm
=countm,counts=counts,count2s=count2s,mucoverage=mcov, sigmacoverage=scov)
   }
```

```
#Simulates TEV 100(1-alpha) CI for lambda^2,
tevsim<-function(n, nruns=5000, mu=0, lam=1, alpha=0.05, d=2*n)
{       theta = 0
            cov <- 0
            low <- 1:nruns
            up <- low
            ucut <- alpha/2
            lcut <- 1-ucut
            lamsq <-lam^2
            for(i in 1:nruns){
                        w<-theta + lam * rexp(n)
                        y<-log(w+1)
                        wn<- 2 * sum(exp(y)-1)
                        low[i] <- wn/qchisq(lcut,df=d)
                        up[i] <- wn/qchisq(ucut, df= d)
                        if(low[i] < lamsq && up[i] > lamsq)
                                       cov <- cov + 1
          }
          cov <- cov / nruns
          slen <- sqrt(n) * mean(up - low)
          list(d = d, cov = cov, slen = slen)
}
```

```
# Simulates Weibull 100(1-alpha) CI for Phi and lambda
"weibsim"<-
function(n = 100, phi = 1, lam = 1, runs = 100, iter = 100)
{  phihat <- 1:runs
lamhat <- phihat
phinew <- 1:runs
     lamnew <- phinew
phio <- 1:runs
lamo <- 1:runs
lnew <- 0
pnew <- 0
    lamold <- 0
lamvold <- 0
phiold <- 0
     phivold <- 0
pcov <- 0
lcov <- 0
pcov2 <- 0
lcov2 <- 0
plow <- 1:runs
pup <- plow
llow <- plow
lup <- plow
AssSDphi <- 0
AssSDlam <- 0
    vec <- 1:4
    for(i in 1:runs)
```

```
{
     # Generating a Weibull R.V
        weib <- (lam * rexp(n))^(1/phi)
        lw <- log(weib)
        tem <- mad(lw, constant = 1)
          phihat[i] <- 0.767049/tem
        ahat <- median(lw) - log(log(2))/phihat[i]
        lamhat[i] <- exp(ahat * phihat[i])


     # Starting values from Olive Robust Estimators: lambda0, Phi0
phiold <- phihat[i]
lamold <- lamhat[i]
    # Calculating MLEs by Iteration
for(j in 1:iter)
{
pnew <-  n/((1/lamold) * sum(weib^phiold * log(weib))-sum(log(weib)))
phivold <- phiold  # = phi[iter-1]
phiold <- pnew
lnew <-  (1/n)*sum(weib^phiold)
lamvold <- lamold  # = lam[iter-1]
lamold <- lnew
}
phio[i] <- phivold
lamo[i] <- lamvold
phinew[i] <- pnew  # MLE
lamnew[i] <- lnew  # MLE
#get CI for phi
```

```
plow[i] <- phinew[i] - 1.96 * .7797 * phinew[i]/sqrt(n)
pup [i] <- phinew[i] + 1.96 * .7797 * phinew[i]/sqrt(n)
if(plow[i] < phi && pup[i] > phi) {pcov <- pcov + 1}
#get CI for lambda
llow[i] <- lamnew[i] - 1.96 * sqrt(1.109*lamnew[i]^2
*(1+.4635*log(lamnew[i])+.5824*(log(lamnew[i]))^2))/sqrt(n)
lup [i] <- lamnew[i] + 1.96 * sqrt(1.109*lamnew[i]^2
*(1+.4635*log(lamnew[i])+.5824*(log(lamnew[i]))^2))/sqrt(n)
if(llow[i] < lam && lup[i] > lam) {lcov <- lcov + 1}
}
 vec[1] <- mean(phinew)
 AsSDphi <- sqrt(.608 * phi/n)
 vec[2] <- sqrt(var(phinew))
 vec[3] <- mean(lamnew)
AsSDlam <- sqrt((.514*lam^2*log(lam)+1.109*lam^2+.608*lam^2
*(log(lam))^2)/n)
 vec[4] <-  sqrt(var(lamnew))
pcov <- pcov / runs
lcov <- lcov / runs
slenp <- sqrt(n) * mean(pup - plow)
slenl <- sqrt(n) * mean(lup - llow)
pconv <- max(abs(phinew - phio))
lconv <- max(abs(lamnew - lamo))
list(phi = phi, phi0 = phihat,phio = phio, phimle = phinew,lamda = lam,
lam0 = lamhat,lamo = lamo,lammle = lamnew, phicov = pcov,slen = slenp,
lamcov = lcov, slenl = slenl, meanphiMLE =vec[1], SDphiMLE = vec[2],
phiSDfromI = AsSDphi, meanlamMLE= vec[3], SDlamMLE = vec[4],
```

```
lamSDfromI=AsSDlam, phiconv = pconv, lamconv = lconv)
   }
```

VITA

Graduate School
Southern Illinois University

Hassan Abuhassan Date of Birth: April 22, 1966

2025 Evergreen Terrace Apt. 6, CARBONDALE, ILLINOIS 62901

Yarmouk University Bachelor of Science, Mathematics, May 1989

The Jordanian University Master of Science, Mathematics, August 1991

Research Paper Title:
Some Transformed Distributions

Major Professor: Dr. David. Olive