Chapter 10

Some Useful Distributions

Definition 10.1. The *population median* is any value MED(Y) such that

$$P(Y \leq \text{MED}(Y)) \geq 0.5 \text{ and } P(Y \geq \text{MED}(Y)) \geq 0.5. \tag{10.1}$$

Definition 10.2. The *population median absolute deviation* is

$$\text{MAD}(Y) = \text{MED}(|Y - \text{MED}(Y)|). \tag{10.2}$$

Definition 10.3. The *gamma function* $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ for $x > 0$.

Some properties of the gamma function follow.
i) $\Gamma(k) = (k-1)!$ for integer $k \geq 1$.
ii) $\Gamma(x+1) = x\ \Gamma(x)$ for $x > 0$.
iii) $\Gamma(x) = (x-1)\ \Gamma(x-1)$ for $x > 1$.
iv) $\Gamma(0.5) = \sqrt{\pi}$.

10.1 The Beta Distribution

If Y has a beta distribution, $Y \sim \text{beta}(\delta, \nu)$, then the probability density function (pdf) of Y is

$$f(y) = \frac{\Gamma(\delta+\nu)}{\Gamma(\delta)\Gamma(\nu)} y^{\delta-1}(1-y)^{\nu-1}$$

where $\delta > 0$, $\nu > 0$ and $0 \le y \le 1$.

$$E(Y) = \frac{\delta}{\delta+\nu}.$$

$$\text{VAR}(Y) = \frac{\delta\nu}{(\delta+\nu)^2(\delta+\nu+1)}.$$

Notice that

$$f(y) = \frac{\Gamma(\delta+\nu)}{\Gamma(\delta)\Gamma(\nu)} I_{[0,1]}(y) \exp[(\delta-1)\log(y) + (\nu-1)\log(1-y)]$$

is a **2P–REF**. Hence $\Theta = (0,\infty) \times (0,\infty)$, $\eta_1 = \delta - 1$, $\eta_2 = \nu - 1$ and $\Omega = (-1,\infty) \times (-1,\infty)$.

If $\delta = 1$, then $W = -\log(1-Y) \sim \text{EXP}(1/\nu)$. Hence $T_n = -\sum \log(1-Y_i) \sim G(n, 1/\nu)$ and if $r > -n$ then T_n^r is the UMVUE of

$$E(T_n^r) = \frac{1}{\nu^r}\frac{\Gamma(r+n)}{\Gamma(n)}.$$

If $\nu = 1$, then $W = -\log(Y) \sim \text{EXP}(1/\delta)$. Hence $T_n = -\sum \log(Y_i) \sim G(n, 1/\delta)$ and and if $r > -n$ then T_n^r is the UMVUE of

$$E(T_n^r) = \frac{1}{\delta^r}\frac{\Gamma(r+n)}{\Gamma(n)}.$$

10.2 The Bernoulli and Binomial Distributions

If Y has a binomial distribution, $Y \sim \text{BIN}(\text{k}, \rho)$, then the probability mass function (pmf) of Y is

$$f(y) = P(Y = y) = \binom{k}{y} \rho^y (1-\rho)^{k-y}$$

for $y = 0, 1, \ldots, k$ where $0 < \rho < 1$.
If $\rho = 0$, $P(Y = 0) = 1 = (1-\rho)^k$ while if $\rho = 1$, $P(Y = k) = 1 = \rho^k$.
The moment generating function

$$m(t) = [(1-\rho) + \rho e^t]^k,$$

and the characteristic function $c(t) = [(1-\rho) + \rho e^{it}]^k$.

$$E(Y) = k\rho.$$

$$\text{VAR}(Y) = k\rho(1-\rho).$$

The Bernoulli (ρ) distribution is the binomial ($k = 1, \rho$) distribution.

The following normal approximation is often used.

$$Y \approx N(k\rho, k\rho(1-\rho))$$

when $k\rho(1-\rho) > 9$. Hence

$$P(Y \le y) \approx \Phi\left(\frac{y + 0.5 - k\rho}{\sqrt{k\rho(1-\rho)}}\right).$$

Also

$$P(Y = y) \approx \frac{1}{\sqrt{k\rho(1-\rho)}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(y-k\rho)^2}{k\rho(1-\rho)}\right).$$

See Johnson, Kotz and Kemp (1992, p. 115). This approximation suggests that $\text{MED}(Y) \approx k\rho$, and $\text{MAD}(Y) \approx 0.674\sqrt{k\rho(1-\rho)}$. Hamza (1995) states that $|E(Y) - \text{MED}(Y)| \le \max(\rho, 1-\rho)$ and shows that

$$|E(Y) - \text{MED}(Y)| \le \log(2).$$

If k is large and $k\rho$ small, then $Y \approx \text{Poisson}(k\rho)$.

If $Y_1, ..., Y_n$ are independent $\text{BIN}(k_i, \rho)$ then $\sum Y_i \sim \text{BIN}(\sum \text{k}_\text{i}, \rho)$.

Notice that

$$f(y) = \binom{k}{y}(1-\rho)^k \exp\left[\log\left(\frac{\rho}{1-\rho}\right)y\right]$$

is a **1P–REF** in ρ if k is known. Thus $\Theta = (0, 1)$,

$$\eta = \log\left(\frac{\rho}{1-\rho}\right)$$

and $\Omega = (-\infty, \infty)$.

Assume that $Y_1, ..., Y_n$ are iid BIN(k, ρ), then

$$T_n = \sum Y_i \sim \text{BIN(nk}, \rho).$$

If k is known, then the likelihood $L(\rho) = c \, \rho^{\sum y_i} (1-\rho)^{nk - \sum y_i}$, and the log likelihood

$$\log(L(\rho)) = d + \log(\rho) \sum y_i + (nk - \sum y_i) \log(1-\rho).$$

Hence

$$\frac{d}{d\rho} \log(L(\rho)) = \frac{\sum y_i}{\rho} + \frac{nk - \sum y_i}{1-\rho}(-1) \stackrel{set}{=} 0,$$

or $(1-\rho)\sum y_i = \rho(nk - \sum y_i)$, or $\sum y_i = \rho nk$ or

$$\hat{\rho} = \sum Y_i/(nk).$$

Notice that

$$\frac{d^2}{d\rho^2} \log(L(\rho)) = \frac{-\sum y_i}{\rho^2} - \frac{nk - \sum y_i}{(1-\rho)^2} < 0$$

if $0 < \sum y_i < nk$. Hence $k\hat{\rho} = \overline{Y}$ is the UMVUE, MLE and MME of $k\rho$ if k is known.

10.3 The Burr Distribution

If Y has a Burr distribution, $Y \sim$ Burr(ϕ, λ), then the pdf of Y is

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

where y, ϕ, and λ are all positive.
The cdf of Y is

$$F(y) = 1 - \exp\left[\frac{-\log(1+y^\phi)}{\lambda}\right] = 1 - (1+y^\phi)^{-1/\lambda} \text{ for } y > 0.$$

MED$(Y) = [e^{\lambda \log(2)} - 1]^{1/\phi}$.
See Patel, Kapadia and Owen (1976, p. 195).
$W = \log(1+Y^\phi)$ is EXP(λ).

Notice that

$$f(y) = \frac{1}{\lambda}\phi y^{\phi-1}\frac{1}{1+y^{\phi}}\exp\left[-\frac{1}{\lambda}\log(1+y^{\phi})\right] I(y>0)$$

is a one parameter exponential family if ϕ is known.

If $Y_1, ..., Y_n$ are iid Burr(λ, ϕ), then

$$T_n = \sum \log(1+Y_i^{\phi}) \sim G(n, \lambda).$$

If ϕ is known, then the likelihood

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

and the log likelihood $\log(L(\lambda)) = d - n\log(\lambda) - \frac{1}{\lambda}\sum \log(1+y_i^{\phi})$. Hence

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

or $\sum \log(1+y_i^{\phi}) = n\lambda$ or

$$\hat{\lambda} = \frac{\sum \log(1+Y_i^{\phi})}{n}.$$

Notice that

$$\frac{d^2}{d\lambda^2}\log(L(\lambda)) = \frac{n}{\lambda^2} - \left.\frac{2\sum \log(1+y_i^{\phi})}{\lambda^2}\right|_{\lambda=\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \frac{2n\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n}{\hat{\lambda}^2} < 0.$$

Thus $\hat{\lambda}$ is the UMVUE and MLE of λ if ϕ is known.

If ϕ is known and $r > -n$, then T_n^r is the UMVUE of

$$E(T_n^r) = \lambda^r\frac{\Gamma(r+n)}{\Gamma(n)}.$$

10.4 The Cauchy Distribution

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

$$f(y) = \frac{\sigma}{\pi}\frac{1}{\sigma^2+(y-\mu)^2} = \frac{1}{\pi\sigma[1+(\frac{y-\mu}{\sigma})^2]}$$

where y and μ are real numbers and $\sigma > 0$.
The cumulative distribution function (cdf) of Y is

$$F(y) = \frac{1}{\pi}[\arctan(\frac{y-\mu}{\sigma}) + \pi/2].$$

See Ferguson (1967, p. 102).
This family is a location–scale family that is symmetric about μ. The moments of Y do not exist, but the chf of Y is

$$c(t) = \exp(it\mu - |t|\sigma).$$

$\text{MED}(Y) = \mu$, the upper quartile $= \mu + \sigma$, and the lower quartile $= \mu - \sigma$.
$\text{MAD}(Y) = F^{-1}(3/4) - \text{MED}(Y) = \sigma$.
If $Y_1, ..., Y_n$ are independent $C(\mu_i, \sigma_i)$, then

$$\sum_{i=1}^n a_i Y_i \sim C(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n |a_i| \sigma_i).$$

In particular, if $Y_1, ..., Y_n$ are iid $C(\mu, \sigma)$, then $\overline{Y} \sim C(\mu, \sigma)$.

10.5 The Chi Distribution

If Y has a chi distribution (also called a p–dimensional Rayleigh distribution), $Y \sim \text{chi}(\text{p}, \sigma)$, then the pdf of Y is

$$f(y) = \frac{y^{p-1} e^{\frac{-1}{2\sigma^2} y^2}}{\sigma^p 2^{\frac{p}{2}-1} \Gamma(p/2)}$$

where $y \geq 0$ and $\sigma, p > 0$. This is a scale family if p is known.

$$E(Y) = \sigma\sqrt{2}\frac{\Gamma(\frac{1+p}{2})}{\Gamma(p/2)}.$$

$$\text{VAR}(Y) = 2\sigma^2 \left[\frac{\Gamma(\frac{2+p}{2})}{\Gamma(p/2)} - \left(\frac{\Gamma(\frac{1+p}{2})}{\Gamma(p/2)} \right)^2 \right],$$

and

$$E(Y^r) = 2^{r/2}\sigma^r \frac{\Gamma(\frac{r+p}{2})}{\Gamma(p/2)}$$

for $r > -p$.
The mode is at $\sigma\sqrt{p-1}$ for $p \geq 1$. See Cohen and Whitten (1988, ch. 10).
Note that $W = Y^2 \sim G(p/2, 2\sigma^2)$.
If $p = 1$, then Y has a half normal distribution, $Y \sim \text{HN}(0, \sigma^2)$.
If $p = 2$, then Y has a Rayleigh distribution, $Y \sim \text{R}(0, \sigma)$.
If $p = 3$, then Y has a Maxwell–Boltzmann distribution (also known as a Boltzmann distribution or a Maxwell distribution), $Y \sim \text{MB}\ (0, \sigma)$.
If p is an integer and $Y \sim \text{chi}(p, 1)$, then $Y^2 \sim \chi^2_p$.

Since

$$f(y) = \frac{1}{2^{\frac{p}{2}-1}\Gamma(p/2)\sigma^p} I(y > 0) \exp[(p-1)\log(y) - \frac{1}{2\sigma^2}y^2],$$

this family appears to be a 2P–REF. Notice that $\Theta = (0, \infty) \times (0, \infty)$, $\eta_1 = p - 1$, $\eta_2 = -1/(2\sigma^2)$, and $\Omega = (-1, \infty) \times (-\infty, 0)$.

If p is known then

$$f(y) = \frac{y^{p-1}}{2^{\frac{p}{2}-1}\Gamma(p/2)} I(y > 0) \frac{1}{\sigma^p} \exp\left[\frac{-1}{2\sigma^2}y^2\right]$$

appears to be a 1P–REF.

If $Y_1, ..., Y_n$ are iid $\text{chi}(p, \sigma)$, then

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

If p is known, then the likelihood

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

and the log likelihood

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

Hence

$$\frac{d}{d(\sigma^2)}\log(\sigma^2) = \frac{-np}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum y_i^2 \stackrel{set}{=} 0,$$

or $\sum y_i^2 = np\sigma^2$ or

$$\hat{\sigma}^2 = \frac{\sum Y_i^2}{np}.$$

Notice that

$$\frac{d^2}{d(\sigma^2)^2}\log(L(\sigma^2)) = \frac{np}{2(\sigma^2)^2} - \frac{\sum y_i^2}{(\sigma^2)^3}\bigg|_{\sigma^2=\hat\sigma^2} = \frac{np}{2(\hat\sigma^2)^2} - \frac{np\hat\sigma}{(\hat\sigma^2)^3}\frac{2}{2} = \frac{-np}{2(\hat\sigma^2)^2} < 0.$$

Thus $\hat\sigma^2$ is the UMVUE and MLE of σ^2 when p is known.

If p is known and $r > -np/2$, then T_n^r is the UMVUE of

$$E(T_n^r) = \frac{2^r\sigma^{2r}\Gamma(r+np/2)}{\Gamma(np/2)}.$$

10.6 The Chi–square Distribution

If Y has a chi–square distribution, $Y \sim \chi^2_p$, then the pdf of Y is

$$f(y) = \frac{y^{\frac{p}{2}-1}e^{-\frac{y}{2}}}{2^{\frac{p}{2}}\Gamma(\frac{p}{2})}$$

where $y \geq 0$ and p is a positive integer.
The mgf of Y is

$$m(t) = \left(\frac{1}{1-2t}\right)^{p/2} = (1-2t)^{-p/2}$$

for $t < 1/2$. The chf

$$c(t) = \left(\frac{1}{1-i2t}\right)^{p/2}.$$

$E(Y) = p$.
$\text{VAR}(Y) = 2p$.
Since Y is gamma $G(\nu = p/2, \lambda = 2)$,

$$E(Y^r) = \frac{2^r\Gamma(r+p/2)}{\Gamma(p/2)},\ r > -p/2.$$

$\text{MED}(Y) \approx p - 2/3$. See Pratt (1968, p. 1470) for more terms in the expansion of $\text{MED}(Y)$.
Empirically,

$$\text{MAD}(Y) \approx \frac{\sqrt{2p}}{1.483}(1-\frac{2}{9p})^2 \approx 0.9536\sqrt{p}.$$

There are several normal approximations for this distribution. The Wilson–Hilferty approximation is

$$\left(\frac{Y}{p}\right)^{\frac{1}{3}} \approx N(1 - \frac{2}{9p}, \frac{2}{9p}).$$

See Bowman and Shenton (1992, p. 6). This approximation gives

$$P(Y \le x) \approx \Phi[((\frac{x}{p})^{1/3} - 1 + 2/9p)\sqrt{9p/2}],$$

and

$$\chi^2_{p,\alpha} \approx p(z_\alpha\sqrt{\frac{2}{9p}} + 1 - \frac{2}{9p})^3$$

where z_α is the standard normal percentile, $\alpha = \Phi(z_\alpha)$. The last approximation is good if $p > -1.24\log(\alpha)$. See Kennedy and Gentle (1980, p. 118).

This family is a one parameter exponential family, but is not a REF since the set of integers does not contain an open interval.

10.7 The Double Exponential Distribution

If Y has a double exponential distribution (or Laplace distribution), $Y \sim DE(\theta, \lambda)$, then the pdf of Y is

$$f(y) = \frac{1}{2\lambda}\exp\left(-\frac{|y-\theta|}{\lambda}\right)$$

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

$$F(y) = 0.5\exp\left(\frac{y-\theta}{\lambda}\right) \quad \text{if} \quad y \le \theta,$$

and

$$F(y) = 1 - 0.5\exp\left(\frac{-(y-\theta)}{\lambda}\right) \quad \text{if} \quad y \ge \theta.$$

This family is a location–scale family which is symmetric about θ.
The mgf

$$m(t) = \exp(\theta t)/(1 - \lambda^2 t^2)$$

for $|t| < 1/\lambda$,
and the chf $c(t) = \exp(\theta i t)/(1 + \lambda^2 t^2)$.

$E(Y) = \theta$, and
MED$(Y) = \theta$.
VAR$(Y) = 2\lambda^2$, and
MAD$(Y) = \log(2)\lambda \approx 0.693\lambda$.
Hence $\lambda = \text{MAD}(Y)/\log(2) \approx 1.443\text{MAD}(Y)$.
To see that MAD$(Y) = \lambda\log(2)$, note that $F(\theta + \lambda\log(2)) = 1 - 0.25 = 0.75$.

The maximum likelihood estimators are $\hat{\theta}_{MLE} = \text{MED}(n)$ and

$$\hat{\lambda}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \text{MED}(n)|.$$

A $100(1-\alpha)\%$ confidence interval (CI) for λ is

$$\left(\frac{2\sum_{i=1}^{n}|Y_i - \text{MED}(n)|}{\chi^2_{2n-1,1-\frac{\alpha}{2}}}, \frac{2\sum_{i=1}^{n}|Y_i - \text{MED}(n)|}{\chi^2_{2n-1,\frac{\alpha}{2}}}\right),$$

and a $100(1-\alpha)\%$ CI for θ is

$$\left(\text{MED}(n) \pm \frac{z_{1-\alpha/2}\sum_{i=1}^{n}|Y_i - \text{MED}(n)|}{n\sqrt{n - z^2_{1-\alpha/2}}}\right)$$

where $\chi^2_{p,\alpha}$ and z_α are the α percentiles of the χ^2_p and standard normal distributions, respectively. See Patel, Kapadia and Owen (1976, p. 194).
$W = |Y - \theta| \sim \text{EXP}(\lambda)$.

Notice that

$$f(y) = \frac{1}{2\lambda}\exp\left[\frac{-1}{\lambda}|y - \theta|\right]$$

is a one parameter exponential family in λ if θ is known.

If $Y_1, ..., Y_n$ are iid $DE(\theta, \lambda)$ then

$$T_n = \sum |Y_i - \theta| \sim G(n, \lambda).$$

If θ is known, then the likelihood

$$L(\lambda) = c\frac{1}{\lambda^n}\exp\left[\frac{-1}{\lambda}\sum|y_i - \theta|\right],$$

and the log likelihood

$$\log(L(\lambda)) = d - n\log(\lambda) - \frac{1}{\lambda}\sum|y_i - \theta|.$$

Hence
$$\frac{d}{d\lambda} \log(L(\lambda)) = \frac{-n}{\lambda} + \frac{1}{\lambda^2} \sum |y_i - \theta| \stackrel{set}{=} 0$$
or $\sum |y_i - \theta| = n\lambda$ or
$$\hat{\lambda} = \frac{\sum |Y_i - \theta|}{n}.$$
Notice that
$$\frac{d^2}{d\lambda^2} \log(L(\lambda)) = \frac{n}{\lambda^2} - \frac{2 \sum |y_i - \theta|}{\lambda^3} \Bigg|_{\lambda=\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \frac{2n\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n}{\hat{\lambda}^2} < 0.$$

Thus $\hat{\lambda}$ is the UMVUE and MLE of λ if θ is known.

10.8 The Exponential Distribution

If Y has an exponential distribution, $Y \sim \text{EXP}(\lambda)$, then the pdf of Y is
$$f(y) = \frac{1}{\lambda} \exp\left(-\frac{y}{\lambda}\right) I(y \geq 0)$$
where $\lambda > 0$.
The cdf of Y is
$$F(y) = 1 - \exp(-y/\lambda), \ \ y \geq 0.$$
This distribution is a scale family.
The mgf
$$m(t) = 1/(1 - \lambda t)$$
for $t < 1/\lambda$, and
the chf $c(t) = 1/(1 - i\lambda t)$.
$E(Y) = \lambda$,
and $\text{VAR}(Y) = \lambda^2$.
$W = 2\lambda Y \sim \chi^2_2$.
Since Y is gamma $G(\nu = 1, \lambda)$, $E(Y^r) = \lambda\Gamma(r + 1)$ for $r > -1$.
$\text{MED}(Y) = \log(2)\lambda$ and
$\text{MAD}(Y) \approx \lambda/2.0781$ since it can be shown that
$$\exp(\text{MAD}(Y)/\lambda) = 1 + \exp(-\text{MAD}(Y)/\lambda).$$

Hence 2.0781 MAD(Y) $\approx \lambda$.
The classical estimator is $\hat{\lambda} = \overline{Y}_n$ and the $100(1-\alpha)$% CI for $E(Y) = \lambda$ is

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

where $P(Y \le \chi^2_{2n,\frac{\alpha}{2}}) = \alpha/2$ if Y is χ^2_{2n}. See Patel, Kapadia and Owen (1976, p. 188).

Notice that

$$f(y) = \frac{1}{\lambda} I(y > 0) \exp\left[\frac{-1}{\lambda} y\right]$$

is a **1P–REF**. Hence $\Theta = (0, \infty)$, $\eta = -1/\lambda$ and $\Omega = (-\infty, 0)$.

Suppose that $Y_1, ..., Y_n$ are iid EXP(λ), then

$$T_n = \sum Y_i \sim G(n, \lambda).$$

The likelihood

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

and the log likelihood

$$\log(L(\lambda)) = -n \log(\lambda) - \frac{1}{\lambda} \sum y_i.$$

Hence

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

or $\sum y_i = n\lambda$ or

$$\hat{\lambda} = \overline{Y}.$$

Since

$$\frac{d^2}{d\lambda^2} \log(L(\lambda)) = \frac{n}{\lambda^2} - \left.\frac{2\sum y_i}{\lambda^3}\right|_{\lambda=\hat{\lambda}} = \frac{n}{\hat{\lambda}^2} - \frac{2n\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n}{\hat{\lambda}^2} < 0,$$

the $\hat{\lambda}$ is the UMVUE, MLE and MME of λ.

If $r > -n$, then T_n^r is the UMVUE of

$$E(T_n^r) = \frac{\lambda^r \Gamma(r+n)}{\Gamma(n)}.$$

10.9 The Two Parameter Exponential Distribution

If Y has a 2 parameter exponential distribution, $Y \sim \text{EXP}(\theta, \lambda)$ then the pdf of Y is

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

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

$$F(y) = 1 - \exp[-(y-\theta)/\lambda)], \ y \geq \theta.$$

This family is an asymmetric location-scale family.
The mgf

$$m(t) = \exp(t\theta)/(1 - \lambda t)$$

for $t < 1/\lambda$, and
the chf $c(t) = \exp(it\theta)/(1 - i\lambda t)$.
$E(Y) = \theta + \lambda$,
and $\text{VAR}(Y) = \lambda^2$.

$$\text{MED}(Y) = \theta + \lambda \log(2)$$

and

$$\text{MAD}(Y) \approx \lambda/2.0781.$$

Hence $\theta \approx \text{MED}(Y) - 2.0781 \log(2)\text{MAD}(Y)$. See Rousseeuw and Croux (1993) for similar results. Note that $2.0781 \log(2) \approx 1.44$.

To see that $2.0781\text{MAD}(Y) \approx \lambda$, note that

$$0.5 = \int_{\theta+\lambda\log(2)-\text{MAD}}^{\theta+\lambda\log(2)+\text{MAD}} \frac{1}{\lambda} \exp(-(y-\theta)/\lambda)dy$$

$$= 0.5[-e^{-\text{MAD}/\lambda} + e^{\text{MAD}/\lambda}]$$

assuming $\lambda \log(2) > \text{MAD}$. Plug in $\text{MAD} = \lambda/2.0781$ to get the result.

If θ is known, then

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

is a 1P–REF in λ. Notice that $Y - \theta \sim EXP(\lambda)$. Let

$$\hat{\lambda} = \frac{\sum (Y_i - \theta)}{n}.$$

Then $\hat{\lambda}$ is the UMVUE and MLE of λ if θ is known.

If $Y_1, ..., Y_n$ are iid EXP(θ, λ), then the likelihood

$$L(\theta, \lambda) = \frac{1}{\lambda^n} \exp\left[\frac{-1}{\lambda} \sum (y_i - \theta)\right] I(y_{(1)} \geq \theta),$$

and the log likelihood

$$\log(L(\lambda)) = [-n \log(\lambda) - \frac{1}{\lambda} \sum (y_i - \theta)] I(y_{(1)} \geq \theta).$$

For any fixed $\lambda > 0$, the log likelihood is maximized by maximizing θ. Hence $\hat{\theta} = Y_{(1)}$, and the profile log likelihood is

$$\log(L(\lambda | y_{(1)})) = -n \log(\lambda) - \frac{1}{\lambda} \sum (y_i - y_{(1)})$$

is maximized by $\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} (y_i - y_{(1)})$. Hence the MLE

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

10.10 The F Distribution

If Y has an F distribution, $Y \sim F(\nu_1, \nu_2)$, then the pdf of Y is

$$f(y) = \frac{\Gamma(\frac{\nu_1 + \nu_2}{2})}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} \frac{y^{(\nu_1 - 2)/2}}{\left(1 + (\frac{\nu_1}{\nu_2}) y\right)^{(\nu_1 + \nu_2)/2}}$$

where $y > 0$ and ν_1 and ν_2 are positive integers.

$$E(Y) = \frac{\nu_2}{\nu_2 - 2}, \quad \nu_2 > 2$$

and

$$\text{VAR}(Y) = 2\left(\frac{\nu_2}{\nu_2 - 2}\right)^2 \frac{(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 4)}, \quad \nu_2 > 4.$$

$$E(Y^r) = \frac{\Gamma(\frac{\nu_1+2r}{2})\Gamma(\frac{\nu_2-2r}{2})}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \left(\frac{\nu_2}{\nu_1}\right)^r, \quad r < \nu_2/2.$$

Suppose that X_1 and X_2 are independent where $X_1 \sim \chi^2_{\nu_1}$ and $X_2 \sim \chi^2_{\nu_2}$. Then

$$W = \frac{(X_1/\nu_1)}{(X_2/\nu_2)} \sim F(\nu_1, \nu_2).$$

If $W \sim t_\nu$, then $Y = W^2 \sim F(1, \nu)$.

10.11 The Gamma Distribution

If Y has a gamma distribution, $Y \sim G(\nu, \lambda)$, then the pdf of Y is

$$f(y) = \frac{y^{\nu-1} e^{-y/\lambda}}{\lambda^\nu \Gamma(\nu)}$$

where ν, λ, and y are positive.
The mgf of Y is

$$m(t) = \left(\frac{1/\lambda}{\frac{1}{\lambda} - t}\right)^\nu = \left(\frac{1}{1 - \lambda t}\right)^\nu$$

for $t < 1/\lambda$. The chf

$$c(t) = \left(\frac{1}{1 - i\lambda t}\right)^\nu.$$

$E(Y) = \nu\lambda$.
$\text{VAR}(Y) = \nu\lambda^2$.

$$E(Y^r) = \frac{\lambda^r \Gamma(r + \nu)}{\Gamma(\nu)} \text{ if } r > -\nu. \tag{10.3}$$

Chen and Rubin (1986) show that $\lambda(\nu - 1/3) < \text{MED}(Y) < \lambda\nu = E(Y)$. Empirically, for $\nu > 3/2$,

$$\text{MED}(Y) \approx \lambda(\nu - 1/3),$$

and

$$\text{MAD}(Y) \approx \frac{\lambda\sqrt{\nu}}{1.483}.$$

This family is a scale family for fixed ν, so if Y is $G(\nu, \lambda)$ then cY is $G(\nu, c\lambda)$ for $c > 0$. If W is EXP(λ) then W is $G(1, \lambda)$. If W is χ^2_p, then W is $G(p/2, 2)$.

Some classical estimates are given next. Let

$$w = \log \left[\frac{\overline{y}_n}{\text{geometric mean}(n)} \right]$$

where geometric mean$(n) = (y_1 y_2 \ldots y_n)^{1/n}$. Then Thom's estimate (Johnson and Kotz 1970a, p. 188) is

$$\hat{\nu} \approx \frac{0.25(1 + \sqrt{1 + 4w/3}\ )}{w}.$$

Also

$$\hat{\nu}_{MLE} \approx \frac{0.5000876 + 0.1648852w - 0.0544274w^2}{w}$$

for $0 < w \leq 0.5772$, and

$$\hat{\nu}_{MLE} \approx \frac{8.898919 + 9.059950w + 0.9775374w^2}{w(17.79728 + 11.968477w + w^2)}$$

for $0.5772 < w \leq 17$. If $W > 17$ then estimation is much more difficult, but a rough approximation is $\hat{\nu} \approx 1/w$ for $w > 17$. See Bowman and Shenton (1988, p. 46) and Greenwood and Durand (1960). Finally, $\hat{\lambda} = \overline{Y}_n/\hat{\nu}$. Notice that $\hat{\beta}$ may not be very good if $\hat{\nu} < 1/17$.

Several normal approximations are available. The Wilson–Hilferty approximation says that for $\nu > 0.5$,

$$Y^{1/3} \approx N\left((\nu\lambda)^{1/3}(1 - \frac{1}{9\nu}), (\nu\lambda)^{2/3}\frac{1}{9\nu} \right).$$

Hence if Y is $G(\nu, \lambda)$ and

$$\alpha = P[Y \leq G_\alpha],$$

then

$$G_\alpha \approx \nu\lambda \left[z_\alpha \sqrt{\frac{1}{9\nu}} + 1 - \frac{1}{9\nu} \right]^3$$

where z_α is the standard normal percentile, $\alpha = \Phi(z_\alpha)$. Bowman and Shenton (1988, p. 101) include higher order terms.

Notice that

$$f(y) = \frac{1}{\lambda^\nu \Gamma(\nu)} I(y > 0) \exp\left[\frac{-1}{\lambda} y + (\nu - 1)\log(y)\right]$$

is a **2P–REF**. Hence $\Theta = (0, \infty) \times (0, \infty)$, $\eta_1 = -1/\lambda$, $\eta_2 = \nu - 1$ and $\Omega = (-\infty, 0) \times (-1, \infty)$.

If $Y_1, ..., Y_n$ are independent $G(\nu_i, \lambda)$ then $\sum Y_i \sim G(\sum \nu_i, \lambda)$.

If $Y_1, ..., Y_n$ are iid $G(\nu, \lambda)$, then

$$T_n = \sum Y_i \sim G(n\nu, \lambda).$$

Since

$$f(y) = \frac{1}{\Gamma(\nu)} \exp[(\nu - 1)\log(y)] I(y > 0) \frac{1}{\lambda^\nu} \exp\left[\frac{-1}{\lambda} y\right],$$

Y is a 1P–REF when ν is known.

If ν is known, then the likelihood

$$L(\beta) = c \, \frac{1}{\lambda^{n\nu}} \exp\left[\frac{-1}{\lambda} \sum y_i\right].$$

The log likelihood

$$\log(L(\lambda)) = d - n\nu \log(\lambda) - \frac{1}{\lambda} \sum y_i.$$

Hence

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

or $\sum y_i = n\nu\lambda$ or

$$\hat{\lambda} = \overline{Y}/\nu.$$

Notice that

$$\frac{d^2}{d\lambda^2} \log(L(\lambda)) = \frac{n\nu}{\lambda^2} - \left.\frac{2\sum y_i}{\lambda^3}\right|_{\lambda = \hat{\lambda}} = \frac{n\nu}{\hat{\lambda}^2} - \frac{2n\nu\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n\nu}{\hat{\lambda}^2} < 0.$$

Thus $\overline{Y}$ is the UMVUE, MLE and MME of $\nu\lambda$ if ν is known.

10.12 The Generalized Gamma Distribution

If Y has a generalized gamma distribution, $Y \sim GG(\nu, \lambda, \phi)$, then the pdf of Y is

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

where ν, λ, ϕ and y are positive.
This family is a scale family with scale parameter λ if ϕ and ν are known.

$$E(Y^r) = \frac{\lambda^r \Gamma(\nu) + \frac{1}{\phi}}{\Gamma(\nu)} \quad \text{if} \quad r > -\phi\nu. \tag{10.4}$$

If ϕ and ν are known, then

$$f(y) = \frac{\phi y^{\phi\nu-1}}{\Gamma(\nu)} I(y > 0) \frac{1}{\lambda^{\phi\nu}} \exp\left[\frac{-1}{\lambda^{\phi}} y^{\phi}\right],$$

which is a one parameter exponential family.

Notice that $W = Y^{\phi} \sim G(\nu, \lambda^{\phi})$. If $Y_1, ..., Y_n$ are iid $GG((\nu, \lambda, \phi)$ where ϕ and ν are known, then $T_n = \sum_{i=1}^n Y_i^{\phi} \sim G(n\nu, \lambda^{\phi})$, and T_n^r is the UMVUE of

$$E(T_n^r) = \lambda^{\phi r} \frac{\Gamma(r + n\nu)}{\Gamma(n\nu)}$$

for $r > -n\nu$.

10.13 The Generalized Negative Binomial Distribution

If Y has a generalized negative binomial distribution, $Y \sim GNB(\mu, \kappa)$, then the pmf of Y is

$$f(y) = P(Y = y) = \frac{\Gamma(y + \kappa)}{\Gamma(\kappa)\Gamma(y + 1)} \left(\frac{\kappa}{\mu + \kappa}\right)^{\kappa} \left(1 - \frac{\kappa}{\mu + \kappa}\right)^{y}$$

for $y = 0, 1, 2, ...$ where $\mu > 0$ and $\kappa > 0$. This distribution is a generalization of the negative binomial (κ, ρ) distribution with $\rho = \kappa/(\mu + \kappa)$ and $\kappa > 0$ is an unknown real parameter rather than a known integer.

The mgf is

$$m(t) = \left[\frac{\kappa}{\kappa + \mu(1 - e^t)}\right]^\kappa$$

for $t < -\log(\mu/(\mu + \kappa))$.
$E(Y) = \mu$ and
VAR$(Y) = \mu + \mu^2/\kappa$.

If $Y_1, ..., Y_n$ are iid GNB(μ, κ), then $\sum_{i=1}^n Y_i \sim GNB(n\mu, n\kappa)$.

When κ is known, this distribution is a **1P–REF**. If $Y_1, ..., Y_n$ are iid GNB(μ, κ) where κ is known, then $\hat{\mu} = \overline{Y}$ is the MLE, UMVUE and MME of μ.

10.14 The Geometric Distribution

If Y has a geometric distribution, $Y \sim$ geom(ρ) then the pmf of Y is

$$f(y) = P(Y = y) = \rho(1 - \rho)^y$$

for $y = 0, 1, 2, ...$ and $0 < \rho < 1$.

The cdf for Y is $F(y) = 1 - (1 - \rho)^{\lfloor y+1 \rfloor}$ for $y \geq 0$ and $F(y) = 0$ for $y < 0$. Here $\lfloor y \rfloor$ is the greatest integer function, eg, $\lfloor 7.7 \rfloor = 7$.

$E(Y) = (1 - \rho)/\rho$.

VAR$(Y) = (1 - \rho)/\rho^2$.

$Y \sim NB(1, \rho)$.

Hence the mgf of Y is

$$m(t) = \frac{\rho}{1 - (1 - \rho)e^t}$$

for $t < -\log(1 - \rho)$.

Notice that

$$f(y) = \rho \exp[\log(1 - \rho)y]$$

is a **1P–REF**. Hence $\Theta = (0, 1)$, $\eta = \log(1 - \rho)$ and $\Omega = (-\infty, 0)$.

If $Y_1, ..., Y_n$ are iid geom(ρ), then

$$T_n = \sum Y_i \sim \text{NB}(\text{n}, \rho).$$

The likelihood

$$L(\rho) = \rho^n \exp[\log(1 - \rho) \sum y_i],$$

and the log likelihood

$$\log(L(\rho)) = n\log(\rho) + \log(1-\rho)\sum y_i.$$

Hence

$$\frac{d}{d\rho}\log(L(\rho)) = \frac{n}{\rho} - \frac{1}{1-\rho}\sum y_i \stackrel{set}{=} 0$$

or $n(1-\rho)/\rho = \sum y_i$ or $n - n\rho - \rho\sum y_i = 0$ or

$$\hat{\rho} = \frac{n}{n + \sum Y_i}.$$

Notice that

$$\frac{d^2}{d\rho^2}\log(L(\rho)) = \frac{-n}{\rho^2} - \frac{\sum y_i}{(1-\rho)^2} < 0.$$

Thus $\hat{\rho}$ is the MLE of ρ.

The UMVUE, MLE and MME of $(1-\rho)/\rho$ is $\overline{Y}$.

10.15 The Half Cauchy Distribution

If Y has a half Cauchy distribution, $Y \sim \text{HC}(\mu, \sigma)$, then the pdf of Y is

$$f(y) = \frac{2}{\pi\sigma[1 + (\frac{y-\mu}{\sigma})^2]}$$

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

$$F(y) = \frac{2}{\pi}\arctan(\frac{y-\mu}{\sigma})$$

for $y \geq \mu$ and is 0, otherwise. This distribution is a right skewed location-scale family.

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

MAD$(Y) = 0.73205\sigma$.

10.16 The Half Logistic Distribution

If Y has a half logistic distribution, $Y \sim \text{HL}(\mu, \sigma)$, then the pdf of y is

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

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

$$F(y) = \frac{\exp[(y-\mu)/\sigma] \ - \ 1}{1 + \exp[(y-\mu)/\sigma)]}$$

for $y \geq \mu$ and 0 otherwise. This family is a right skewed location–scale family.

MED(Y) $= \mu + \log(3)\sigma$.

MAD(Y) $= 0.67346\sigma$.

10.17 The Half Normal Distribution

If Y has a half normal distribution, $Y \sim \text{HN}(\mu, \sigma^2)$, 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. 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.

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

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

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.

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

MAD(Y) $= 0.3990916\sigma$.

Notice that

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

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(μ, σ^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

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

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

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

Notice that

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

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

Thus $\hat{\sigma}^2$ is the UMVUE and MLE of σ^2 if μ is known.

If $r > -n/2$ and if μ is known, then T_n^r is the UMVUE of

$$E(T_n^r) = 2^r \sigma^{2r} \Gamma(r + n/2)/\Gamma(n/2).$$

10.18 The Hypergeometric Distribution

If Y has a hypergeometric distribution, $Y \sim \mathrm{HG(C, N - C, n)}$, then the data set contains N objects of two types. There are C objects of the first type (that you wish to count) and $N - C$ objects of the second type. Suppose that n objects are selected at random without replacement from the N objects.

Then Y counts the number of the n selected objects that were of the first type. The pmf of Y is

$$f(y) = P(Y = y) = \frac{\binom{C}{y}\binom{N-C}{n-y}}{\binom{N}{n}}$$

where the integer y satisfies $\max(0, n - N + C) \le y \le \min(n, C)$. The right inequality is true since if n objects are selected, then the number of objects of the first type must be less than or equal to both n and C. The first inequality holds since $n - Y$ counts the number of objects of second type. Hence $n - Y \le N - C$.

Let $p = C/N$. Then

$$E(Y) = \frac{nC}{N} = np$$

and

$$\text{VAR}(Y) = \frac{nC(N-C)}{N^2}\frac{N-n}{N-1} = np(1-p)\frac{N-n}{N-1}.$$

If n is small compared to both C and $N - C$ then $Y \approx \text{BIN(n, p)}$. If n is large but n is small compared to both C and $N-C$ then $Y \approx N(np, np(1-p))$.

10.19 The Inverse Gaussian Distribution

If Y has an inverse Gaussian distribution, $Y \sim \text{IG}(\theta, \lambda)$, then the pdf of Y is

$$f(y) = \sqrt{\frac{\lambda}{2\pi y^3}} \exp\left[\frac{-\lambda(y-\theta)^2}{2\theta^2 y}\right]$$

where $y, \theta, \lambda > 0$.
The mgf is

$$m(t) = \exp\left[\frac{\lambda}{\theta}\left(1 - \sqrt{1 - \frac{2\theta^2 t}{\lambda}}\right)\right]$$

for $t < \lambda/(2\theta^2)$. See Datta (2005) and Schwarz and Samanta (1991) for additional properties.

The chf is

$$\phi(t) = \exp\left[\frac{\lambda}{\theta}\left(1 - \sqrt{1 - \frac{2\theta^2 it}{\lambda}}\right)\right].$$

$E(Y) = \theta$ and

$$\text{VAR}(Y) = \frac{\theta^3}{\lambda}.$$

Notice that

$$f(y) = \sqrt{\frac{\lambda}{2\pi}} e^{\lambda/\theta} \sqrt{\frac{1}{y^3}} I(y > 0) \exp\left[\frac{-\lambda}{2\theta^2} y - \frac{\lambda}{2} \frac{1}{y}\right]$$

is a two parameter exponential family.

If $Y_1, ..., Y_n$ are iid $IG(\theta, \lambda)$, then

$$\sum_{i=1}^{n} Y_i \sim IG(n\theta, n^2\lambda) \quad \text{and} \quad \overline{Y} \sim IG(\theta, n\lambda).$$

If λ is known, then the likelihood

$$L(\theta) = c\ e^{n\lambda/\theta} \exp[\frac{-\lambda}{2\theta^2} \sum y_i],$$

and the log likelihood

$$\log(L(\theta)) = d + \frac{n\lambda}{\theta} - \frac{\lambda}{2\theta^2} \sum y_i.$$

Hence

$$\frac{d}{d\theta} \log(L(\theta)) = \frac{-n\lambda}{\theta^2} + \frac{\lambda}{\theta^3} \sum y_i \stackrel{set}{=} 0,$$

or $\sum y_i = n\theta$ or

$$\hat{\theta} = \overline{Y}.$$

Notice that

$$\frac{d^2}{d\theta^2} \log(L(\theta)) = \frac{2n\lambda}{\theta^3} - \frac{3\lambda \sum y_i}{\theta^4}\Bigg|_{\theta=\hat{\theta}} = \frac{2n\lambda}{\hat{\theta}^3} - \frac{3n\lambda\hat{\theta}}{\hat{\theta}^4} = \frac{-n\lambda}{\hat{\theta}^3} < 0.$$

Thus $\overline{Y}$ is the UMVUE, MLE and MME of θ if λ is known.

If θ is known, then the likelihood

$$L(\lambda) = c\ \lambda^{n/2} \exp\left[\frac{-\lambda}{2\theta^2} \sum \frac{(y_i - \theta)^2}{y_i}\right],$$

and the log likelihood

$$\log(L(\lambda)) = d + \frac{n}{2}\log(\lambda) - \frac{\lambda}{2\theta^2}\sum \frac{(y_i-\theta)^2}{y_i}.$$

Hence

$$\frac{d}{d\lambda}\log(L(\lambda)) = \frac{n}{2\lambda} - \frac{1}{2\theta^2}\sum \frac{(y_i-\theta)^2}{y_i} \overset{set}{=} 0$$

or

$$\hat{\lambda} = \frac{n\theta^2}{\sum \frac{(y_i-\theta)^2}{y_i}}.$$

Notice that

$$\frac{d^2}{d\lambda^2}\log(L(\lambda)) = \frac{-n}{2\lambda^2} < 0.$$

Thus $\hat{\lambda}$ is the MLE of λ if θ is known.

Another parameterization of the inverse Gaussian distribution takes $\theta = \sqrt{\lambda/\psi}$ so that

$$f(y) = \sqrt{\frac{\lambda}{2\pi}}e^{\sqrt{\lambda\psi}}\sqrt{\frac{1}{y^3}}I[y>0]\exp\left[\frac{-\psi}{2}y - \frac{\lambda}{2}\frac{1}{y}\right],$$

where $\lambda > 0$ and $\psi \geq 0$. Here $\Theta = (0,\infty)\times[0,\infty)$, $\eta_1 = -\psi/2$, $\eta_2 = \lambda/2$ and $\Omega = (-\infty, 0] \times (-\infty, 0)$. Since Ω is not an open set, this is a **2 parameter full exponential family that is not regular**. If ψ is known then Y is a 1P–REF, but if λ is known the Y is a one parameter full exponential family. When $\psi = 0$, Y has a one sided stable distribution with index 1/2. See Barndorff–Nielsen (1978, p. 117).

10.20 The Inverted Gamma Distribution

If Y has an inverted gamma distribution, $Y \sim INVG(\nu, \lambda)$, then the pdf of Y is

$$f(y) = \frac{1}{y^{\nu+1}\Gamma(\nu)}I(y>0)\frac{1}{\lambda^\nu}\exp\left(\frac{-1}{\lambda}\frac{1}{y}\right)$$

where λ, ν and y are all positive. It can be shown that $W = 1/Y \sim G(\nu, \lambda)$. This family is a scale family with scale parameter $\tau = 1/\lambda$ if ν is known.

If ν is known, this family is a 1 parameter exponential family. If $Y_1, ..., Y_n$ are iid INVG(ν, λ) and ν is known, then $T_n = \sum_{i=1}^n \frac{1}{Y_i} \sim G(n\nu, \lambda)$ and T_n^r is the UMVUE of

$$\lambda^r \frac{\Gamma(r + n\nu)}{\Gamma(n\nu)}$$

for $r > -n\nu$.

10.21 The Largest Extreme Value Distribution

If Y has a largest extreme value distribution (or Gumbel distribution), $Y \sim LEV(\theta, \sigma)$, 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 $\sigma > 0$. The cdf of Y is

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

This family is an asymmetric location–scale family with a mode at θ.
The mgf

$$m(t) = \exp(t\theta)\Gamma(1 - \sigma t)$$

for $|t| < 1/\sigma$.
$E(Y) \approx \theta + 0.57721\sigma$, and
$\text{VAR}(Y) = \sigma^2\pi^2/6 \approx 1.64493\sigma^2$.

$$\text{MED}(Y) = \theta - \sigma \log(\log(2)) \approx \theta + 0.36651\sigma$$

and

$$\text{MAD}(Y) \approx 0.767049\sigma.$$

$W = \exp(-(Y - \theta)/\sigma) \sim \text{EXP}(1)$.

Notice that

$$f(y) = \frac{1}{\sigma} e^{\theta/\sigma} e^{-y/\sigma} \exp\left[-e^{\theta/\sigma} e^{-y/\sigma}\right]$$

is a one parameter exponential family in θ if σ is known.

If $Y_1, ..., Y_n$ are iid LEV$(\theta, \sigma)$where σ is known, then the likelihood

$$L(\sigma) = c \; e^{n\theta/\sigma} \exp[-e^{\theta/\sigma} \sum e^{-y_i/\sigma}],$$

and the log likelihood

$$\log(L(\theta)) = d + \frac{n\theta}{\sigma} - e^{\theta/\sigma} \sum e^{-y_i/\sigma}.$$

Hence

$$\frac{d}{d\theta} \log(L(\theta)) = \frac{n}{\sigma} - e^{\theta/\sigma} \frac{1}{\sigma} \sum e^{-y_i/\sigma} \stackrel{set}{=} 0,$$

or

$$e^{\theta/\sigma} \sum e^{-y_i/\sigma} = n,$$

or

$$e^{\theta/\sigma} = \frac{n}{\sum e^{-y_i/\sigma}},$$

or

$$\hat{\theta} = \log\left(\frac{n}{\sum e^{-Y_i/\sigma}}\right).$$

Since

$$\frac{d^2}{d\theta^2} \log(L(\theta)) = \frac{-1}{\sigma^2} e^{\theta/\sigma} \sum e^{-y_i/\sigma} < 0,$$

$\hat{\theta}$ is the MLE of θ.

10.22 The Logarithmic Distribution

If Y has a logarithmic distribution, then the pmf of Y is

$$f(y) = P(Y = y) = \frac{-1}{\log(1-\theta)} \frac{\theta^y}{y}$$

for $y = 1, 2, ...$ and $0 < \theta < 1$.

The mgf

$$m(t) = \frac{\log(1 - \theta e^t)}{\log(1-\theta)}$$

for $t < -\log(\theta)$.

$$E(Y) = \frac{-1}{\log(1-\theta)} \frac{\theta}{1-\theta}.$$

Notice that
$$f(y) = \frac{-1}{\log(1-\theta)} \frac{1}{y} \exp(\log(\theta) y)$$
is a **1P–REF**. Hence $\Theta = (0,1)$, $\eta = \log(\theta)$ and $\Omega = (-\infty, 0)$.

If $Y_1, ..., Y_n$ are iid logarithmic (θ), then $\overline{Y}$ is the UMVUE of $E(Y)$.

10.23 The Logistic Distribution

If Y has a logistic distribution, $Y \sim L(\mu, \sigma)$, then the pdf of y is
$$f(y) = \frac{\exp\left(-(y-\mu)/\sigma\right)}{\sigma[1+\exp\left(-(y-\mu)/\sigma\right)]^2}$$
where $\sigma > 0$ and y and μ are real.
The cdf of Y is
$$F(y) = \frac{1}{1+\exp\left(-(y-\mu)/\sigma\right)} = \frac{\exp\left((y-\mu)/\sigma\right)}{1+\exp\left((y-\mu)/\sigma\right)}.$$
This family is a symmetric location–scale family.
The mgf of Y is $m(t) = \pi\sigma t e^{\mu t} \csc(\pi\sigma t)$ for $|t| < 1/\sigma$, and
the chf is $c(t) = \pi i\sigma t e^{i\mu t} \csc(\pi i\sigma t)$ where $csc(t)$ is the cosecant of t.
$E(Y) = \mu$, and
$\text{MED}(Y) = \mu$.
$\text{VAR}(Y) = \sigma^2\pi^2/3$, and
$\text{MAD}(Y) = \log(3)\sigma \approx 1.0986\ \sigma$.
Hence $\sigma = \text{MAD}(Y)/\log(3)$.
The estimators $\hat{\mu} = \overline{Y}_n$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \overline{Y}_n)^2$ are sometimes used.
Note that if
$$q = F_{L(0,1)}(c) = \frac{e^c}{1+e^c} \quad \text{then} \quad c = \log(\frac{q}{1-q}).$$
Taking $q = .9995$ gives $c = \log(1999) \approx 7.6$.
To see that $\text{MAD}(Y) = \log(3)\sigma$, note that $F(\mu + \log(3)\sigma) = 0.75$,
$F(\mu - \log(3)\sigma) = 0.25$, and $0.75 = \exp\left(\log(3)\right)/(1+\exp(\log(3)))$.

10.24 The Log-Cauchy Distribution

If Y has a log–Cauchy distribution , $Y \sim LC(\mu, \sigma)$, then the pdf of Y is

$$f(y) = \frac{1}{\pi \sigma y [1 + (\frac{\log(y) - \mu}{\sigma})^2]}$$

where $y > 0$, $\sigma > 0$ and μ is a real number. This family is a scale family with scale parameter $\tau = e^{\mu}$ if σ^2 is known. It can be shown that $W = \log(Y)$ has a Cauchy(μ, σ) distribution.

10.25 The Log-Logistic Distribution

If Y has a log–logistic distribution, $Y \sim LL(\phi, \tau)$, then the pdf of Y is

$$f(y) = \frac{\phi \tau (\phi y)^{\tau - 1}}{[1 + (\phi y)^{\tau}]^2}$$

where $y > 0$, $\phi > 0$ and $\tau > 0$. The cdf of Y is

$$F(y) = 1 - \frac{1}{1 + (\phi y)^{\tau}}$$

for $y > 0$. This family is a scale family with scale parameter $\sigma = \phi^{-1}$ if τ is known.

MED$(Y) = 1/\phi$.

It can be shown that $W = \log(Y)$ has a logistic$(\mu = -\log(\phi), \sigma = 1/\tau)$ distribution. Hence $\phi = e^{-\mu}$ and $\tau = 1/\sigma$. Kalbfleisch and Prentice (1980, p. 27-28) suggest that the log-logistic distribution is a competitor of the lognormal distribution.

10.26 The Lognormal Distribution

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

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

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

$$F(y) = \Phi\left(\frac{\log(y) - \mu}{\sigma}\right) \quad \text{for} \quad y > 0$$

where $\Phi(y)$ is the standard normal N(0,1) cdf.
This family is a scale family with scale parameter $\tau = e^{\mu}$ if σ^2 is known.

$$E(Y) = \exp(\mu + \sigma^2/2)$$

and

$$\text{VAR}(Y) = \exp(\sigma^2)(\exp(\sigma^2) - 1)\exp(2\mu).$$

For any r,

$$E(Y^r) = \exp(r\mu + r^2\sigma^2/2).$$

MED$(Y) = \exp(\mu)$ and
$\exp(\mu)[1 - \exp(-0.6744\sigma)] \leq \text{MAD}(Y) \leq \exp(\mu)[1 + \exp(0.6744\sigma)]$.

Notice that

$$f(y) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}\exp(\frac{-\mu^2}{2\sigma^2})\frac{1}{y}I(y \geq 0)\exp\left[\frac{-1}{2\sigma^2}(\log(y))^2 + \frac{\mu}{\sigma^2}\log(y)\right]$$

is a **2P–REF**. Hence $\Theta = (-\infty, \infty) \times (0, \infty)$, $\eta_1 = -1/(2\sigma^2)$, $\eta_2 = \mu/\sigma^2$ and $\Omega = (-\infty, 0) \times (-\infty, \infty)$.

Note that $W = \log(Y) \sim N(\mu, \sigma^2)$.

Notice that

$$f(y) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}\frac{1}{y}I(y \geq 0)\exp\left[\frac{-1}{2\sigma^2}(\log(y) - \mu)^2\right]$$

is a 1P–REF if μ is known,.

If $Y_1, ..., Y_n$ are iid LN(μ, σ^2) where μ is known, then the likelihood

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

and the log likelihood

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

Hence

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

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

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

Since

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

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

$\hat{\sigma}^2$ is the UMVUE and MLE of σ^2 if μ is known.

Since $T_n = \sum[\log(Y_i) - \mu]^2 \sim G(n/2, 2\sigma^2)$, if μ is known and $r > -n/2$ then T_n^r is UMVUE of

$$E(T_n^r) = 2^r \sigma^{2r} \frac{\Gamma(r + n/2)}{\Gamma(n/2)}.$$

If σ^2 is known,

$$f(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \frac{1}{y} I(y \ge 0) \exp(\frac{-1}{2\sigma^2}(\log(y))^2) \exp(\frac{-\mu^2}{2\sigma^2}) \exp\left[\frac{\mu}{\sigma^2} \log(y)\right]$$

is a 1P–REF.

If $Y_1, ..., Y_n$ are iid LN(μ, σ^2), where σ^2 is known, then the likelihood

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

and the log likelihood

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

Hence

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

or $\sum \log(y_i) = n\mu$ or

$$\hat{\mu} = \frac{\sum \log(Y_i)}{n}.$$

Note that

$$\frac{d^2}{d\mu^2} \log(L(\mu)) = \frac{-n}{\sigma^2} < 0.$$

Since $T_n = \sum \log(Y_i) \sim N(n\mu, n\sigma^2)$, $\hat{\mu}$ is the UMVUE and MLE of μ if σ^2 is known.

When neither μ nor σ are known, the log likelihood

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

Let $w_i = \log(y_i)$ then the log likelihood is

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

which is equivalent to the normal $N(\mu, \sigma^2)$ log likelihood. Hence the MLE

$$(\hat{\mu}, \hat{\sigma}) = \left(\frac{1}{n}\sum_{i=1}^{n} W_i, \sqrt{\frac{1}{n}\sum_{i=1}^{n}(W_i - \overline{W})^2} \right).$$

Hence inference for μ and σ is simple. Use the fact that $W_i = \log(Y_i) \sim N(\mu, \sigma^2)$ and then perform the corresponding normal based inference on the W_i. For example, a the classical $(1-\alpha)100\%$ CI for μ when σ is unknown is

$$(\overline{W}_n - t_{n-1,1-\frac{\alpha}{2}}\frac{S_W}{\sqrt{n}}, \overline{W}_n + t_{n-1,1-\frac{\alpha}{2}}\frac{S_W}{\sqrt{n}})$$

where

$$S_W = \frac{n}{n-1}\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(W_i - \overline{W})^2},$$

and $P(t \leq t_{n-1,1-\frac{\alpha}{2}}) = 1 - \alpha/2$ when t is from a t distribution with $n-1$ degrees of freedom. Compare Meeker and Escobar (1998, p. 175).

10.27 The Maxwell-Boltzmann Distribution

If Y has a Maxwell–Boltzmann distribution, $Y \sim MB(\mu, \sigma)$, then the pdf of Y is

$$f(y) = \frac{\sqrt{2}(y-\mu)^2 e^{\frac{-1}{2\sigma^2}(y-\mu)^2}}{\sigma^3\sqrt{\pi}}$$

where μ is real, $y \geq \mu$ and $\sigma > 0$. This is a location–scale family.

$$E(Y) = \mu + \sigma\sqrt{2}\frac{1}{\Gamma(3/2)}.$$

$$\text{VAR}(Y) = 2\sigma^2 \left[\frac{\Gamma(\frac{5}{2})}{\Gamma(3/2)} - \left(\frac{1}{\Gamma(3/2)} \right)^2 \right].$$

MED(Y) $= \mu + 1.5381722\sigma$ and MAD(Y) $= 0.460244\sigma$.
This distribution a one parameter exponential family when μ is known.
Note that $W = (Y - \mu)^2 \sim G(3/2, 2\sigma^2)$.
If $Z \sim MB(0, \sigma)$, then $Z \sim \text{chi}(\text{p} = 3, \sigma)$, and

$$E(Z^r) = 2^{r/2}\sigma^r \frac{\Gamma(\frac{r+3}{2})}{\Gamma(3/2)}$$

for $r > -3$.
The mode of Z is at $\sigma\sqrt{2}$.

10.28 The Negative Binomial Distribution

If Y has a negative binomial distribution (also called the Pascal distribution), $Y \sim \text{NB}(\text{r}, \rho)$, then the pmf of Y is

$$f(y) = P(Y = y) = \binom{r + y - 1}{y} \rho^r (1 - \rho)^y$$

for $y = 0, 1, \ldots$ where $0 < \rho < 1$.
The moment generating function

$$m(t) = \left[\frac{\rho}{1 - (1 - \rho)e^t} \right]^r$$

for $t < -\log(1 - \rho)$.
$E(Y) = r(1 - \rho)/\rho$, and

$$\text{VAR}(Y) = \frac{r(1 - \rho)}{\rho^2}.$$

Notice that

$$f(y) = \rho^r \binom{r + y - 1}{y} \exp[\log(1 - \rho)y]$$

is a **1P–REF** in ρ for known r. Thus $\Theta = (0, 1)$, $\eta = \log(1 - \rho)$ and $\Omega = (-\infty, 0)$.

If $Y_1, ..., Y_n$ are independent NB(r_i, ρ), then

$$\sum Y_i \sim \text{NB}(\sum \text{r}_\text{i}, \rho).$$

If $Y_1, ..., Y_n$ are iid $NB(r, \rho)$, then

$$T_n = \sum Y_i \sim NB(nr, \rho).$$

If r is known, then the likelihood

$$L(p) = c\ \rho^{nr} \exp[\log(1-\rho) \sum y_i],$$

and the log likelihood

$$\log(L(\rho)) = d + nr\log(\rho) + \log(1-\rho)\sum y_i.$$

Hence

$$\frac{d}{d\rho}\log(L(\rho)) = \frac{nr}{\rho} - \frac{1}{1-\rho}\sum y_i \stackrel{set}{=} 0,$$

or

$$\frac{1-\rho}{\rho} nr = \sum y_i,$$

or $nr - \rho nr - \rho \sum y_i = 0$ or

$$\hat{\rho} = \frac{nr}{nr + \sum Y_i}.$$

Notice that

$$\frac{d^2}{d\rho^2}\log(L(\rho)) = \frac{-nr}{\rho^2} - \frac{1}{(1-\rho)^2}\sum y_i < 0.$$

Thus $\hat{\rho}$ is the MLE of p if r is known.

Notice that $\overline{Y}$ is the UMVUE, MLE and MME of $r(1-\rho)/\rho$ if r is known.

10.29 The Normal Distribution

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

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

where $\sigma > 0$ and μ and y are real.
Let $\Phi(y)$ denote the standard normal cdf. Recall that $\Phi(y) = 1 - \Phi(-y)$. The cdf $F(y)$ of Y does not have a closed form, but

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

and

$$\Phi(y) \approx 0.5(1 + \sqrt{1 - \exp(-2y^2/\pi)}\ )$$

for $y \geq 0$. See Johnson and Kotz (1970a, p. 57).
The moment generating function is

$$m(t) = \exp(t\mu + t^2\sigma^2/2).$$

The characteristic function is $c(t) = \exp(it\mu - t^2\sigma^2/2)$.
$E(Y) = \mu$ and
$\text{VAR}(Y) = \sigma^2$.

$$E[|Y-\mu|^r] = \sigma^r \ \frac{2^{r/2}\Gamma((r+1)/2)}{\sqrt{\pi}} \quad \text{for} \ \ r > -1.$$

If $k \geq 2$ is an integer, then $E(Y^k) = (k-1)\sigma^2 E(Y^{k-2}) + \mu E(Y^{k-1})$. See Stein (1981) and Casella and Berger (1990, p. 187).
$\text{MED}(Y) = \mu$ and

$$\text{MAD}(Y) = \Phi^{-1}(0.75)\sigma \approx 0.6745\sigma.$$

Hence $\sigma = [\Phi^{-1}(0.75)]^{-1}\text{MAD}(Y) \approx 1.483\text{MAD}(Y)$.
This family is a location–scale family which is symmetric about μ.
Suggested estimators are

$$\overline{Y}_n = \hat{\mu} = \frac{1}{n}\sum_{i=1}^n Y_i \text{ and } S^2 = S_Y^2 = \frac{1}{n-1}\sum_{i=1}^n (Y_i - \overline{Y}_n)^2.$$

The classical $(1-\alpha)100\%$ CI for μ when σ is unknown is

$$(\overline{Y}_n - t_{n-1,1-\frac{\alpha}{2}}\frac{S_Y}{\sqrt{n}}, \overline{Y}_n + t_{n-1,1-\frac{\alpha}{2}}\frac{S_Y}{\sqrt{n}})$$

where $P(t \le t_{n-1,1-\frac{\alpha}{2}}) = 1 - \alpha/2$ when t is from a t distribution with $n-1$ degrees of freedom.
If $\alpha = \Phi(z_\alpha)$, then

$$z_\alpha \approx m - \frac{c_o + c_1 m + c_2 m^2}{1 + d_1 m + d_2 m^2 + d_3 m^3}$$

where

$$m = [-2\log(1-\alpha)]^{1/2},$$

$c_0 = 2.515517$, $c_1 = 0.802853$, $c_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, $d_3 = 0.001308$, and $0.5 \le \alpha$. For $0 < \alpha < 0.5$,

$$z_\alpha = -z_{1-\alpha}.$$

See Kennedy and Gentle (1980, p. 95).
To see that $\text{MAD}(Y) = \Phi^{-1}(0.75)\sigma$, note that $3/4 = F(\mu + \text{MAD})$ since Y is symmetric about μ. However,

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

and

$$\frac{3}{4} = \Phi\left(\frac{\mu + \Phi^{-1}(3/4)\sigma - \mu}{\sigma}\right).$$

So $\mu + \text{MAD} = \mu + \Phi^{-1}(3/4)\sigma$. Cancel μ from both sides to get the result.

Notice that

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

is a **2P–REF**. Hence $\Theta = (0,\infty) \times (-\infty,\infty)$, $\eta_1 = -1/(2\sigma^2)$, $\eta_2 = \mu/\sigma^2$ and $\Omega = (-\infty, 0) \times (-\infty, \infty)$.

If σ^2 is known,

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

is a 1P–REF. Also the likelihood

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

and the log likelihood

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

Hence

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

or $n\mu = \sum y_i$, or

$$\hat{\mu} = \overline{Y}.$$

Since

$$\frac{d^2}{d\mu^2}\log(L(\mu)) = \frac{-n}{\sigma^2} < 0$$

and since $T_n = \sum Y_i \sim N(n\mu, n\sigma^2)$, $\overline{Y}$ is the UMVUE, MLE and MME of μ if σ^2 is known.

If μ is known,

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

is a 1P–REF. Also 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

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

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

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

Note that

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

$$= \frac{-n}{2(\hat{\sigma}^2)^2} < 0.$$

Since $T_n = \sum(Y_i - \mu)^2 \sim G(n/2, 2\sigma^2)$, $\hat{\sigma}^2$ is the UMVUE and MLE of σ^2 if μ is known.

Note that if μ is known and $r > -n/2$, then T_n^r is the UMVUE of

$$E(T_n^r) = 2^r \sigma^{2r} \frac{\Gamma(r + n/2)}{\Gamma(n/2)}.$$

10.30 The One Sided Stable Distribution

If Y has a one sided stable distribution (with index 1/2), $Y \sim OSS(\sigma)$, then the pdf of Y is

$$f(y) = \frac{1}{\sqrt{2\pi y^3}} \sqrt{\sigma} \exp\left(\frac{-\sigma}{2} \frac{1}{y}\right)$$

for $y > 0$ and $\sigma > 0$. This distribution is a scale family with scale parameter σ and a **1P–REF**. When $\sigma = 1$, $Y \sim \text{INVG}(\nu = 1/2, \lambda = 2)$ where INVG stands for inverted gamma. This family is a special case of the inverse Gaussian IG distribution. It can be shown that $W = 1/Y \sim G(1/2, 2/\sigma)$. This distribution is even more outlier prone than the Cauchy distribution. See Feller (1971, p. 52) and Lehmann (1999, p. 76). For applications see Besbeas and Morgan (2004).

If $Y_1, ..., Y_n$ are iid OSS(σ) then $T_n = \sum_{i=1}^n \frac{1}{Y_i} \sim G(n/2, 2/\sigma)$. Hence T_n/n is the UMVUE (and MLE) of $1/\sigma$ and T_n^r is the UMVUE of

$$\frac{1}{\sigma^r} \frac{2^r \Gamma(r + n/2)}{\Gamma(n/2)}$$

for $r > -n/2$.

10.31 The Pareto Distribution

If Y has a Pareto distribution, $Y \sim \text{PAR}(\sigma, \lambda)$, 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 mode is at $Y = \sigma$.
The cdf of Y is $F(y) = 1 - (\sigma/y)^{1/\lambda}$ for $y > \mu$.
This family is a scale family when λ is fixed.

$$E(Y) = \frac{\sigma}{1-\lambda}$$

for $\lambda < 1$.

$$E(Y^r) = \frac{\sigma^r}{1-r\lambda} \text{ for } r < 1/\lambda.$$

$\text{MED}(Y) = \sigma 2^\lambda$.
$X = \log(Y/\sigma)$ is $\text{EXP}(\lambda)$ and $W = \log(Y)$ is $\text{EXP}(\theta = \log(\sigma), \lambda)$.

Notice that

$$f(y) = \frac{1}{\lambda}\sigma^{1/\lambda} I[y \geq \sigma] \exp\left[-(1+\frac{1}{\lambda})\log(y)\right] =$$

$$\frac{1}{\sigma}\frac{1}{\lambda} I[y \geq \sigma] \exp\left[-(1+\frac{1}{\lambda})\log(y/\sigma)\right]$$

is a one parameter exponential family if σ is known.

If $Y_1, ..., Y_n$ are iid $\text{PAR}(\sigma, \lambda)$ then

$$T_n = \sum \log(Y_i/\sigma) \sim G(n, \lambda).$$

If σ is known, then the likelihood

$$L(\lambda) = c\frac{1}{\lambda^n} \exp\left[-(1+\frac{1}{\lambda})\sum \log(y_i/\sigma)\right],$$

and the log likelihood

$$\log(L(\lambda)) = d - n\log(\lambda) - (1+\frac{1}{\lambda})\sum \log(y_i/\sigma).$$

Hence

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

or $\sum \log(y_i/\sigma) = n\lambda$ or

$$\hat{\lambda} = \frac{\sum \log(Y_i/\sigma)}{n}.$$

Notice that

$$\frac{d^2}{d\lambda^2}\log(L(\lambda)) = \frac{n}{\lambda^2} - \left.\frac{2\sum \log(y_i/\sigma)}{\lambda^3}\right|_{\lambda=\hat{\lambda}} =$$

$$\frac{n}{\hat{\lambda}^2} - \frac{2n\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n}{\hat{\lambda}^2} < 0.$$

Hence $\hat{\lambda}$ is the UMVUE and MLE of λ if σ is known.

If σ is known and $r > -n$, then T_n^r is the UMVUE of

$$E(T_n^r) = \lambda^r \frac{\Gamma(r+n)}{\Gamma(n)}.$$

If neither σ nor λ are known, notice that

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

Hence the likelihood

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

and the log likelihood is

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

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

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

which is equivalent to 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

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

10.32 The Poisson Distribution

If Y has a Poisson distribution, $Y \sim \text{POIS}(\theta)$, then the pmf of Y is

$$f(y) = P(Y = y) = \frac{e^{-\theta}\theta^y}{y!}$$

for $y = 0, 1, \ldots,$ where $\theta > 0$.
If $\theta = 0$, $P(Y = 0) = 1 = e^{-\theta}$.
The mgf of Y is

$$m(t) = \exp(\theta(e^t - 1)),$$

and the chf of Y is
$c(t) = \exp(\theta(e^{it} - 1))$.
$E(Y) = \theta$, and Chen and Rubin (1986) and Adell and Jodrá (2005) show that $-1 < \text{MED}(Y) - E(Y) < 1/3$.
$\text{VAR}(Y) = \theta$.
The classical estimator of θ is $\hat{\theta} = \overline{Y}_n$.
The approximations $Y \approx N(\theta, \theta)$ and $2\sqrt{Y} \approx N(2\sqrt{\theta}, 1)$ are sometimes used.

Notice that

$$f(y) = e^{-\theta}\frac{1}{y!}\exp[\log(\theta)y]$$

is a **1P–REF**. Thus $\Theta = (0, \infty)$, $\eta = \log(\theta)$ and $\Omega = (-\infty, \infty)$.

If $Y_1, ..., Y_n$ are independent $\text{POIS}(\theta_i)$ then $\sum Y_i \sim \text{POIS}(\sum \theta_{\text{i}})$.

If $Y_1, ..., Y_n$ are iid $POIS(\theta)$ then

$$T_n = \sum Y_i \sim \text{POIS}(\text{n}\theta).$$

The likelihood

$$L(\theta) = c\ e^{-n\theta}\exp[\log(\theta)\sum y_i],$$

and the log likelihood

$$\log(L(\theta)) = d - n\theta + \log(\theta)\sum y_i.$$

Hence

$$\frac{d}{d\theta}\log(L(\theta)) = -n + \frac{1}{\theta}\sum y_i \stackrel{set}{=} 0,$$

or $\sum y_i = n\theta$, or

$$\hat{\theta} = \overline{Y}.$$

Notice that
$$\frac{d^2}{d\theta^2} \log(L(\theta)) = \frac{-\sum y_i}{\theta^2} < 0$$
unless $\sum y_i = 0$.

Hence $\overline{Y}$ is the UMVUE and MLE of θ.

10.33 The Power Distribution

If Y has a power distribution, $Y \sim \text{POW}(\lambda)$, then the pdf of Y is
$$f(y) = \frac{1}{\lambda} y^{\frac{1}{\lambda}-1},$$
where $\lambda > 0$ and $0 \le y \le 1$.
The cdf of Y is $F(y) = y^{1/\lambda}$ for $0 \le y \le 1$.
$\text{MED}(Y) = (1/2)^\lambda$.
$W = -\log(Y)$ is $\text{EXP}(\lambda)$. Notice that $Y \sim \text{beta}(\delta = 1/\lambda, \nu = 1)$.

Notice that
$$f(y) = \frac{1}{\lambda} I_{[0,1]}(y) \exp\left[(\frac{1}{\lambda} - 1) \log(y) \right]$$
$$= \frac{1}{\lambda} I_{[0,1]}(y) \exp\left[(1 - \frac{1}{\lambda})(-\log(y)) \right]$$
is a **1P–REF**. Thus $\Theta = (0, \infty)$, $\eta = 1 - 1/\lambda$ and $\Omega = (-\infty, 1)$.

If $Y_1, ..., Y_n$ are iid $POW(\lambda)$, then
$$T_n = -\sum \log(Y_i) \sim G(n, \lambda).$$
The likelihood
$$L(\lambda) = \frac{1}{\lambda^n} \exp\left[(\frac{1}{\lambda} - 1) \sum \log(y_i) \right],$$
and the log likelihood
$$\log(L(\lambda)) = -n \log(\lambda) + (\frac{1}{\lambda} - 1) \sum \log(y_i).$$
Hence
$$\frac{d}{d\lambda} \log(L(\lambda)) = \frac{-n}{\lambda} - \frac{\sum \log(y_i)}{\lambda^2} \stackrel{set}{=} 0,$$

or $-\sum \log(y_i) = n\lambda$, or
$$\hat{\lambda} = \frac{-\sum \log(Y_i)}{n}.$$

Notice that
$$\frac{d^2}{d\lambda^2} \log(L(\lambda)) = \frac{n}{\lambda^2} - \left.\frac{2\sum \log(y_i)}{\lambda^3}\right|_{\lambda=\hat{\lambda}}$$
$$= \frac{n}{\hat{\lambda}^2} + \frac{2n\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n}{\hat{\lambda}^2} < 0.$$

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

If $r > -n$, then T_n^r is the UMVUE of
$$E(T_n^r) = \lambda^r \frac{\Gamma(r+n)}{\Gamma(n)}.$$

10.34 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 $y \geq \mu$. See Cohen and Whitten (1988, Ch. 10). This is an asymmetric location–scale family.

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.
$$E(Y) = \mu + \sigma\sqrt{\pi/2} \approx \mu + 1.253314\sigma.$$
$$\text{VAR}(Y) = \sigma^2(4-\pi)/2 \approx 0.429204\sigma^2.$$
$\text{MED}(Y) = \mu + \sigma\sqrt{\log(4)} \approx \mu + 1.17741\sigma.$

Hence $\mu \approx \text{MED}(Y) - 2.6255\text{MAD}(Y)$ and $\sigma \approx 2.230\text{MAD}(Y)$.

Let $\sigma D = \text{MAD}(Y)$. If $\mu = 0$, and $\sigma = 1$, then
$$0.5 = \exp[-0.5(\sqrt{\log(4)} - D)^2] - \exp[-0.5(\sqrt{\log(4)} + D)^2].$$

Hence $D \approx 0.448453$ and $\text{MAD}(Y) \approx 0.448453\sigma$.
It can be shown that $W = (Y - \mu)^2 \sim \text{EXP}(2\sigma^2)$.

Other parameterizations for the Rayleigh distribution are possible.

Note that

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

appears to be a 1P–REF if μ is known.

If $Y_1, ..., Y_n$ are iid $R(\mu, \sigma)$, then

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

If μ is known, then the likelihood

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

and the log likelihood

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

Hence

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

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

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

Notice that

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

$$\frac{n}{(\hat{\sigma}^2)^2} - \frac{2n\hat{\sigma}^2}{(\hat{\sigma}^2)^3} = \frac{-n}{(\hat{\sigma}^2)^2} < 0.$$

Hence $\hat{\sigma}^2$ is the UMVUE and MLE of σ^2 if μ is known.

If μ is known and $r > -n$, then T_n^r is the UMVUE of

$$E(T_n^r) = 2^r \sigma^{2r} \frac{\Gamma(r + n)}{\Gamma(n)}.$$

10.35 The Smallest Extreme Value Distribution

If Y has a smallest extreme value distribution (or log-Weibull 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 $\sigma > 0$.
The cdf of Y is

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

This family is an asymmetric location-scale family with a longer left tail than right.

$E(Y) \approx \theta - 0.57721\sigma$, and
VAR(Y) $= \sigma^2\pi^2/6 \approx 1.64493\sigma^2$.

MED(Y) $= \theta - \sigma \log(\log(2))$.

MAD(Y) $\approx 0.767049\sigma$.

Y is a one parameter exponential family in θ if σ is known.

If Y has a SEV(θ, σ) distribution, then $W = -Y$ has an LEV($-\theta, \sigma$) distribution.

10.36 The Student's t Distribution

If Y has a Student's t distribution, $Y \sim t_p$, then the pdf of Y is

$$f(y) = \frac{\Gamma(\frac{p+1}{2})}{(p\pi)^{1/2}\Gamma(p/2)} (1 + \frac{y^2}{p})^{-(\frac{p+1}{2})}$$

where p is a positive integer and y is real. This family is symmetric about 0. The t_1 distribution is the Cauchy$(0,1)$ distribution. If Z is $N(0,1)$ and is independent of $W \sim \chi^2_p$, then

$$\frac{Z}{(\frac{W}{p})^{1/2}}$$

is t_p.
$E(Y) = 0$ for $p \geq 2$.
MED(Y) $= 0$.

VAR$(Y) = p/(p-2)$ for $p \geq 3$, and
MAD$(Y) = t_{p,0.75}$ where $P(t_p \leq t_{p,0.75}) = 0.75$.
If $\alpha = P(t_p \leq t_{p,\alpha})$, then Cooke, Craven, and Clarke (1982, p. 84) suggest the approximation

$$t_{p,\alpha} \approx \sqrt{p[\exp(\frac{w_\alpha^2}{p}) - 1)]}$$

where

$$w_\alpha = \frac{z_\alpha(8p+3)}{8p+1},$$

z_α is the standard normal cutoff: $\alpha = \Phi(z_\alpha)$, and $0.5 \leq \alpha$. If $0 < \alpha < 0.5$, then

$$t_{p,\alpha} = -t_{p,1-\alpha}.$$

This approximation seems to get better as the degrees of freedom increase.

10.37 The Truncated Extreme Value Distribution

If Y has a truncated extreme value distribution, $Y \sim$ TEV(λ), then the pdf of Y is

$$f(y) = \frac{1}{\lambda} \exp\left(y - \frac{e^y - 1}{\lambda}\right)$$

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

$$F(y) = 1 - \exp\left[\frac{-(e^y - 1)}{\lambda}\right]$$

for $y > 0$.
MED$(Y) = \log(1 + \lambda \log(2))$.
$W = e^Y - 1$ is EXP(λ).

Notice that

$$f(y) = \frac{1}{\lambda} e^y I(y \geq 0) \exp\left[\frac{-1}{\lambda}(e^y - 1)\right]$$

is a **1P–REF**. Hence $\Theta = (0, \infty)$, $\eta = -1/\lambda$ and $\Omega = (-\infty, 0)$.

If $Y_1, ..., Y_n$ are iid TEV(λ), then

$$T_n = \sum (e^{Y_i} - 1) \sim G(n, \lambda).$$

The likelihood

$$L(\lambda) = c \, \frac{1}{\lambda^n} \exp\left[\frac{-1}{\lambda} \sum \log(e^{y_i} - 1)\right],$$

and the log likelihood

$$\log(L(\lambda)) = d - n\log(\lambda) - \frac{1}{\lambda} \sum \log(e^{y_i} - 1).$$

Hence

$$\frac{d}{d\lambda} \log(L(\lambda)) = \frac{-n}{\lambda} + \frac{\sum \log(e^{y_i} - 1)}{\lambda^2} \stackrel{set}{=} 0,$$

or $\sum \log(e^{y_i} - 1) = n\lambda$, or

$$\hat{\lambda} = \frac{-\sum \log(e^{Y_i} - 1)}{n}.$$

Notice that

$$\frac{d^2}{d\lambda^2} \log(L(\lambda)) = \frac{n}{\lambda^2} - \frac{2\sum \log(e^{y_i} - 1)}{\lambda^3}\Bigg|_{\lambda=\hat{\lambda}}$$

$$= \frac{n}{\hat{\lambda}^2} - \frac{2n\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n}{\hat{\lambda}^2} < 0.$$

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

If $r > -n$, then T_n^r is the UMVUE of

$$E(T_n^r) = \lambda^r \frac{\Gamma(r+n)}{\Gamma(n)}.$$

10.38 The Uniform Distribution

If Y has a uniform distribution, $Y \sim U(\theta_1, \theta_2)$, then the pdf of Y is

$$f(y) = \frac{1}{\theta_2 - \theta_1} I(\theta_1 \le y \le \theta_2).$$

The cdf of Y is $F(y) = (y - \theta_1)/(\theta_2 - \theta_1)$ for $\theta_1 \le y \le \theta_2$.
This family is a location-scale family which is symmetric about $(\theta_1 + \theta_2)/2$.
By definition, $m(0) = c(0) = 1$. For $t \neq 0$, the mgf of Y is

$$m(t) = \frac{e^{t\theta_2} - e^{t\theta_1}}{(\theta_2 - \theta_1)t},$$

and the chf of Y is
$$c(t) = \frac{e^{it\theta_2} - e^{it\theta_1}}{(\theta_2 - \theta_1)it}.$$

$E(Y) = (\theta_1 + \theta_2)/2$, and
MED$(Y) = (\theta_1 + \theta_2)/2$.
VAR$(Y) = (\theta_2 - \theta_1)^2/12$, and
MAD$(Y) = (\theta_2 - \theta_1)/4$.
Note that $\theta_1 = \text{MED}(Y) - 2\text{MAD}(Y)$ and $\theta_2 = \text{MED}(Y) + 2\text{MAD}(Y)$.
Some classical estimates are $\hat{\theta}_1 = y_{(1)}$ and $\hat{\theta}_2 = y_{(n)}$.

10.39 The Weibull Distribution

If Y has a Weibull distribution, $Y \sim W(\phi, \lambda)$, then the pdf of Y is
$$f(y) = \frac{\phi}{\lambda} y^{\phi-1} e^{-\frac{y^\phi}{\lambda}}$$
where λ, y, and ϕ are all positive. For fixed ϕ, this is a scale family in $\sigma = \lambda^{1/\phi}$.
The cdf of Y is $F(y) = 1 - \exp(-y^\phi/\lambda)$ for $y > 0$.
$E(Y) = \lambda^{1/\phi}\ \Gamma(1 + 1/\phi)$.
VAR$(Y) = \lambda^{2/\phi}\Gamma(1 + 2/\phi)\ -\ (E(Y))^2$.
$$E(Y^r) = \lambda^{r/\phi}\ \Gamma(1 + \frac{r}{\phi}) \quad \text{for} \quad r > -\phi.$$
MED$(Y) = (\lambda \log(2))^{1/\phi}$.
Note that
$$\lambda = \frac{(\text{MED}(Y))^\phi}{\log(2)}.$$

$W = Y^\phi$ is EXP(λ).

$W = \log(Y)$ has a smallest extreme value SEV$(\theta = \log(\lambda^{1/\phi}), \sigma = 1/\phi)$ distribution.

Notice that
$$f(y) = \frac{\phi}{\lambda} y^{\phi-1} I(y \geq 0) \exp\left[\frac{-1}{\lambda} y^\phi\right]$$
is a one parameter exponential family in λ if ϕ is known.

If $Y_1, ..., Y_n$ are iid $W(\phi, \lambda)$, then

$$T_n = \sum Y_i^\phi \sim G(n, \lambda).$$

If ϕ is known, then the likelihood

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

and the log likelihood

$$\log(L(\lambda)) = d - n\log(\lambda) - \frac{1}{\lambda}\sum y_i^\phi.$$

Hence

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

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

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

Notice that

$$\frac{d^2}{d\lambda^2}\log(L(\lambda)) = \left.\frac{n}{\lambda^2} - \frac{2\sum y_i^\phi}{\lambda^3}\right|_{\lambda=\hat{\lambda}}$$

$$= \frac{n}{\hat{\lambda}^2} - \frac{2n\hat{\lambda}}{\hat{\lambda}^3} = \frac{-n}{\hat{\lambda}^2} < 0.$$

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

If $r > -n$, then T_n^r is the UMVUE of

$$E(T_n^r) = \lambda^r \frac{\Gamma(r+n)}{\Gamma(n)}.$$

10.40 The Zeta Distribution

If Y has a Zeta distribution, $Y \sim Zeta(\nu)$, then the pmf of Y is

$$f(y) = P(Y = y) = \frac{1}{y^\nu \zeta(\nu)}$$

where $\nu > 1$ and $y = 1, 2, 3, ...$ Here the zeta function

$$\zeta(\nu) = \sum_{y=1}^{\infty} \frac{1}{y^\nu}$$

for $\nu > 1$. This distribution is a one parameter exponential family.

$$E(y) = \frac{\zeta(\nu - 1)}{\zeta(\nu)}$$

for $\nu > 2$, and

$$\text{VAR}(Y) = \frac{\zeta(\nu - 2)}{\zeta(\nu)} - \left[\frac{\zeta(\nu - 1)}{\zeta(\nu)}\right]^2$$

for $\nu > 3$.

$$E(Y^r) = \frac{\zeta(\nu - r)}{\zeta(\nu)}$$

for $\nu > r + 1$.

This distribution is sometimes used for count data, especially by linguistics for word frequency. See Lindsey (2004, p. 154).

10.41 Complements

Many of the distribution results used in this chapter came from Johnson and Kotz (1970a,b) and Patel, Kapadia and Owen (1976). Cohen and Whitten (1988), Ferguson (1967), Castillo (1988), Cramér (1946), Kennedy and Gentle (1980), Lehmann (1983), Meeker and Escobar (1998), Bickel and Doksum (1977), DeGroot and Schervish (2001), Hastings and Peacock (1975) and Leemis (1986) also have useful results on distributions. Also see articles in Kotz and Johnson (1982ab, 1983ab, 1985ab, 1986, 1988ab). Often an entire book is devoted to a single distribution, see for example, Bowman and Shenton (1988).