Journal of the Research Institute of Basic Sciences,
Korea Maritime University, Vol. 4, 1994.

Hierarchical and Empirical Bayes Estimation of Exponential Mean Parameter

Sang Jeen Lee* and Choon Il Park**

* Department of Statistics,
Pusan National University, Pusan, Korea
** Department of Applied Mathematics,
Korea Maritime University, Pusan, Korea

1. Introduction

We consider the simultaneous estimation problem of parameters from several independent exponential-type distributions. First, Berger (1980) studied quite extensively with some approach methods. He suggested various improving approach methods of usual estimators on the exponential mean vector under different quadratic losses. and proposed a number of estimators using these approach methods. He also improved an estimator on the natural parameter vector of some exponential type distributions under the squared error loss.

Estimators alternate to Berger's in the gamma case was proposed by Das Gupta (1986). And Ghosh and Parsian (1980), Ghosh, Hwang and Tsui (1984) and Ghosh and Day (1984) generalized Berger's (1980) results.

Let $X_i = (X_{1i}, \ldots, X_{n_i i})$, $i = 1, \ldots, p$ be independent random samples of size n_i from exponential distribution with parameters $\theta = (\theta_1, \ldots, \theta_p)$. Equivalently $Y = (Y_1, \ldots, Y_p)$ where $Y_i = \sum_{j=1}^{n_i} X_{ji}$ is the sufficient statistic having pdf

$$f(y|\theta) = \prod_{i=1}^{p} \frac{\theta_i^{n_i}}{\Gamma(n_i)} y_i^{n_i - 1} \exp(-\theta_i y_i), \quad \forall y_i > 0 \tag{1}$$

where all θ_i's (> 0) is unknown and all sample sizes n_i are given. We want to estimate $\theta = (\theta_1, \ldots, \theta_p)$ by $\delta = (\delta_1, \ldots, \delta_p)$ under the loss

$$L(\theta, \delta) = \sum_{i=1}^{p} [\delta_i \theta_i^{-1} - log(\delta_i \theta_i^{-1}) - 1], \tag{2}$$

and for estimating the mean vector $\theta^{-1} = (\theta_1^{-1}, \ldots, \theta_p^{-1})$, the loss

$$L(\theta^{-1}, \delta) = \sum_{i=1}^{p} [\delta_i \theta_i - log(\delta_i \theta_i) - 1]. \tag{3}$$

These losses are called as entropy losses which correspond to an entropy measure of distance between distributions indexed by θ and δ. An analogous loss was considered in James and Stein (1961) for the estimation of the variance-covariance matrix of a multinormal distribution.

The usual estimator over the vector of means θ^{-1}

$$\begin{aligned} \delta^0(\mathbf{Y}) =& (\frac{Y_1}{n_1}, \ldots, \frac{Y_p}{n_p}) \\ =& (\bar{X}_1, \ldots, \bar{X}_p) \\ =& \bar{\mathbf{X}} \end{aligned}$$

where $\bar{X}_i = Y_i/n_i,\ i = 1, \ldots, p$ is the uniformly minimum variance unbiased estimator (UMVUE) as well as the maximum likelihood estimator (MLE).

Are the above losses invariant ? Yes ! Since θ_i's are scale parameters, we consider the group $G = \{g_c : g_c(x) = cx, c > 0\}$. $\bar{G}$, $\tilde{G}$ and G are isomorphic. For every g in G and a in $\tilde{G}$,

$$L(\{\bar{g}_{\mathbf{c}}(\theta)\}^{-1}, \tilde{g}_{\mathbf{k}}(a)) = \sum_{i=1}^{p}[\bar{g}_{c_i}(\theta_i)\cdot\tilde{g}_{k_i}(a_i) - \log(\bar{g}_{c_i}(\theta_i)\cdot\tilde{g}_{k_i}(a_i)) - 1]$$
$$= \sum_{i=1}^{p}[c_ik_i\theta_ia_i - \log(c_ik_i\theta_ia_i) - 1].$$

If we choose $k_i = \frac{1}{c_i}$, $i = 1, \ldots, p$, it is equal to $L(\theta^{-1}, a)$. Hence there exists $a^* = \tilde{g}_k(a)$, where $k = (k_1, \ldots, k_p)$ and $k_i = \frac{1}{c_i}$, $i = 1, \ldots, p$. Therefore the loss (3) is invariant under $G = \{g_c : g_c(x) = cx, c > 0\}$. Similarly to the loss (2), it is also invariant under the group G.

We will find the best invariant estimator $\delta^I(\mathbf{y})$ the best invariant estimator of θ^{-1}, under the loss (3). But the estimator $\delta^I(\mathbf{y})$ is same to $\delta^0(\mathbf{y})$. If $min_i\{n_i\} > 1$, the best invariant estimator, of θ under the loss (2),

$$\delta^+(Y) = (\frac{n_1 - 1}{Y_1}, \ldots, \frac{n_p - 1}{Y_p})$$

is UMVUE.

Let $R(\theta, \delta) = E_\theta^Y[L(\theta, \delta)]$ denote the risk function of decision rule $\delta(y)$ at a particular value θ. And we say that δ^* uniformly dominates δ, if $R(\theta, \delta^*) \leq R(\theta, \delta)$ for all values of θ with strict inequality for some θ. In this case, we say δ^* is improved from δ.

In univariate case, that is $p = 1$, δ^0 and δ^+ are admissible for θ^{-1} and θ, respectively, from Stein (1959) and Brown (1966). For $p = 2$, Dey et al. (1987) proved the admissibility of $\delta^0(Y)$ for estimating θ^{-1} under the loss

(3), when $\min\{n_1, n_2\} > 4$. They did for $p = 2$ the admissibility of $\delta^+(Y)$ for estimating θ under the loss (2), when $\min\{n_1, n_2\} > 5$. They also proved the inadmissibility of $\delta^0(Y)$ and $\delta^+(Y)$ for estimating θ^{-1} and θ respectively, when $p \geq 3$. They showed the inadmissibility for the generalized Bayes estimator of $(\theta_1^{b_1}, \ldots, \theta_p^{b_p})$ with respect to (possibly improper) prior with pdf $\prod_{i=1}^p \theta^{k_i}$ under the loss

$$\sum_{i=1}^{p} \theta_i^{m_i} [n_i \theta_i^{-b_i} - \log(n_i \theta_i^{-b_i}) - 1] \tag{4}$$

provided certain relationships exist among the m_i's, n_i's, k_i's and b_i's.

Berger (1980) and Das Gupta (1986) improved the best scale invariant estimator under a generalized weighted quadratic loss. Their method is finding the solution δ^*, to

$$\Delta(y_1, \ldots, y_p) < 0$$

where $\Delta(y_1, \ldots, y_p) = R(\theta, \delta^*) - R(\theta, \delta^I)$. componentwise, the estimators are of the form

$$\frac{y_i}{n_i + 1}(1 + \phi(\mathbf{y})).$$

Ghosh and Auer (1983) developed a pseudo-Bayes estimator under squared error loss using an inverse gamma prior distribution, when all sample sizes n_i's are equal. It is determined by adjusting the shrinkng factor of Bayes estimator toward a prior Bayes estimator by a multiple of k whre k is chosen so that it minimizes the risk of new estimator. Albert (81) proposed the estimator of Poisson parameter similarly to above approach.

It will be considered the development of exponential parameter estimators by controlling the shrinking factor of Bayes estimator toward a prior

Bayes estimator under above entropy losses using appropriate prior distribution. We will consider a hierarchical Bayes approach. To model the similarity in size of the parameters $\theta_1, \ldots, \theta_p$, we assume that $\theta_1, \ldots, \theta_p$ are exchangeable. This prior information is represented by letting $\theta_1^{-1}, \ldots, \theta_p^{-1}$ be a random sample from the conjugate prior

$$\pi(\theta_i^{-1}|\alpha,\beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \frac{1}{(\theta_i^{-1})^{\alpha+1}} \exp(\frac{-1}{\beta\theta_i^{-1}}), \quad \alpha,\beta > 0 \tag{5}$$

with hyperparameters α and β. It will be said in section 2 to reparameterize (α,β) into (μ,γ). μ is used as so-called prior Bayes estimator of θ^{-1} and given by

$$\mu = \frac{1}{\alpha\beta}.$$

The parameter μ is the best prior estimate of θ^{-1} in the sense that the prior expected loss is minimized componentwisely at $\delta_i = \mu$.

We will consider the cases when μ is known and when μ is unknown, with interest. This considering is for the goodness of prior selection. That is, it is possible that, through experience, the scientist may believe he or she has a reasonable prior guess for the mean lifetimes in which case μ is considered known. On the other hand, the scientist may believe that the past data has not been gathered accurately enough to make a reasonable guess and thus desire to assume μ to be unknown. If μ is known, the hierarchical model is as follows:

$$\theta^{-1}|\mu,\gamma \sim \text{inverse} - \text{Gamma}(\mu,\gamma)$$

$$\gamma \sim \pi_2$$

where the second stage prior π_2 is noninformative. In contrast, μ is unknown, the hierarchical model is as above with the last expression replaced by

$$(\mu,\gamma) \sim \pi_2$$

where π_2 is now a noninformative prior for the hyperparameters μ and γ. The noninformative prior is used to model the uncertainty of the location of the unknown hyperparameters.

The best invariant estimator will be found and hierarchical Bayes (HB) estimators are developed for μ in section 2. In addition, the empirical Bayes estimators are improved for μ. In section 3, we will compare the two approaches suggesting the preference of the HB approach. The hierarchical Bayes approach is a purely Bayesian approach with a hierarchical prior and the empirical Bayes approach estimates the unknown hyperparameters of the first stage prior from the marginal density using maximum likelihood. In section 4, the μ unknown case is considered. Since the purely HB approach is difficult in computation, it is avoided. However a pseudo-hierarchical approach or a combination of HB and EB approach is employed.

For comparison of estimators in this paper, Dey et al. (1987) is used. We will conclude by briefly outlining how one can develop approximations to estimators.

2. Improvement of HB and EB Estimators

Let us find the best invariant estimator, $\delta^I(\mathbf{y})$, of the exponential scale parameters which situation is specified in previous section, under the entropy loss in above section. For convention, let $\lambda_i = \theta_i^{-1}$, $i = 1, \ldots, p$, that is $\frac{1}{\lambda_i} = \theta_i$. Then, the loss function is as following

$$L(\lambda, a) = \sum_{i=1}^{p}(\frac{a_i}{\lambda_i} - \log(\frac{a_i}{\lambda_i}) - 1)$$

And to convert to location parameter problem, let $\eta_i = \log \lambda_i$, $z_i = \log y_i$ and action $a^* = \log a$ $i = 1, \ldots, p$. Then componentwisely $\delta^{I*}(z_i) =$

$z_i + k$. To find the value of k,

$$L(\eta_i, \delta^{I*}(z_i)) = \exp(z_i + k - \eta_i) - (z_i + k - \eta_i) - 1.$$

Hence, the risk is

$$\begin{aligned} R(\eta_i, \delta^{I*}(Z_i)) =& EL(\eta_i, \delta^{I*}(Z_i)) \\ =& E_0 L(\eta_i, \delta^{I*}(Z_i)) \\ =& E_0(\exp(Z_i + k) - (Z_i + k) - 1) \\ =& e^k E_0(e^{Z_i}) - E_0 Z_i - k - 1. \end{aligned}$$

To minimize this value,

$$\frac{dR}{dk} = e^k E_0 e^{Z_i} - 1 = 0$$

The best invariant estimator in location parameter problem is

$$\delta^{I*}(z_i) = z_i + \log(\frac{1}{E_0 e^{Z_i}})$$

So, the estimator in original problem is

$$\begin{aligned} \delta^I(y_i) =& \exp\{\delta^{I*}(\log y_i)\} \\ =& \frac{y_i}{n_i} \\ =& \bar{x_i}. \end{aligned}$$

In our situation, the best invariant estimator under the entropy loss is equal to UMVU estimator and MLE.

Now, we consider improving problem through HB approach when the best prior estimator μ is known. The prior density of λ is given by

$$\pi(\lambda|\mathbf{y},\alpha,\beta)=\prod_{i=1}^{p}\frac{1}{\beta^{\alpha}\Gamma(\alpha)}\frac{1}{\lambda_i^{\alpha+1}}\exp(\frac{-1}{\beta\lambda_i})$$

Thus, the posterior density of λ given $\mathbf{y},\alpha$ and β is given by

$$\pi(\lambda|\mathbf{y},\alpha,\beta)=\prod_{i=1}^{p}\{\frac{(y_i+\frac{1}{\beta})^{\alpha+n_i}}{\Gamma(\alpha+n_i)}\frac{1}{\lambda_i^{\alpha+n_i+1}}\exp(-(y_i+\beta^{-1})\lambda^{-1})\}, \quad (6)$$

a product of p independent inverse Gamma $(\alpha+n_i, y_i+\beta^{-1})$ distributions for $i = 1,\ldots,p$.

Assume that $\lambda_1,\ldots,\lambda_p$ are exchangeable. For reparameterization of (α,β) to (μ,γ), we find the best prior estimator μ, the prior expected loss with respect to the prior of λ_i, inverse Gamma (α,β), is

$$\begin{aligned}E^{\lambda_i|\alpha,\beta}L(\lambda_i,\delta_i)=&E^{\lambda_i|\alpha,\beta}(\delta_i/\lambda_i-\log(\delta_i/\lambda_i)-1)\\=&\delta_iE^{\lambda_i|\alpha,\beta}\frac{1}{\lambda_i}-\log\delta_i-E^{\lambda_i|\alpha,\beta}\log(\frac{1}{\lambda_i})-1.\end{aligned}$$

For minimizing this,

$$\frac{dEL}{d\delta_i}=E^{\lambda_i|\alpha,\beta}\frac{1}{\lambda_i}-\delta_i^{-1}=0.$$

Therefore, we find the best prior estimator,

$$\begin{aligned}\mu=&\frac{1}{E^{\lambda_i|\alpha,\beta}\frac{1}{\lambda_i}}\\=&\frac{1}{\alpha\beta},\end{aligned} \quad (7)$$

since

$$E^{\lambda_i|\alpha,\beta}(1/\lambda_i) = \int \lambda_i^{-1} \frac{\exp(-1/(\beta\lambda_i))}{\Gamma(\alpha)\beta^\alpha \lambda_i^{\alpha+1}} d\lambda_i$$
$$= \frac{\Gamma(\alpha+1)\beta^{\alpha+1}}{\Gamma(\alpha)\beta^\alpha}$$
$$= \alpha \cdot \beta.$$

The estimate at the first stage of the Bayesian hierarchical analysis, conditional on the values of $\mathbf{y}, \alpha$ and μ, under the above model and the entropy loss, is given by $\hat{\lambda} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_p)$ whose component is

$$\hat{\lambda}_i = \frac{1}{E^{\lambda_i|\mathbf{y}}(1/\lambda_i)}$$
$$= \frac{y_i + \alpha\mu}{\alpha + n_i}$$
$$= \frac{n_i}{\alpha + n_i}\bar{X}_i + \frac{\alpha}{\alpha + n_i}\mu,$$

since

$$E^{\lambda_i|\mathbf{y}}(1/\lambda_i) = \frac{\Gamma(\alpha + n_i + 1)}{\Gamma(\alpha + n_i)}(y_i + \beta^{-1})^{-1}$$
$$= (\alpha + n_i)(y_i + \beta^{-1})^{-1}.$$

When vague prior of λ, $\pi(\lambda) = \prod_{i=1}^p \lambda_i^{-1}$, which can be taken from inverse Gamma with parameters that α approaches 0, is considered, the Bayes estimator is

$$\hat{\lambda}_{iv} = \bar{X}_i,$$

which is from $\hat{\lambda}_i$ by letting α approach 0. It is the best invariant UMVU estimator as well as MLE. We can hence represent the Bayes estimator, $\hat{\lambda}_i$, as a convex combination of the UMVUE $\bar{X}_i$ and the best prior estimator μ.

Consider the limiting value of $\hat{\lambda_i}$ as α approaches ∞.

$$\begin{aligned}\lim_{\alpha\to\infty}\hat{\lambda_i} &= \lim_{\alpha\to\infty}\frac{\mu(n_i+\alpha)}{y_i\mu+\alpha}\\ &=\mu.\end{aligned}$$

The Bayes estimator, $\hat{\lambda_i}$, converges to the best prior estimator, μ, as α approaches ∞.

For more reparameterizarion, let α be $1/\gamma$. Then the Bayes estimator is transformed as

$$\begin{aligned}\hat{\lambda_i} =&E^{\lambda_i|\mathbf{y},\gamma,\mu}(1/\lambda_i)\\ =&\frac{n_i}{\gamma^{-1}+n_i}\bar{X}_i+\frac{\gamma^{-1}}{\gamma^{-1}+n_i})\mu\\ =&\frac{\gamma n_i}{1+\gamma n_i}\bar{X}_i+(1-\frac{\gamma n_i}{1+\gamma n_i})\mu.\end{aligned}$$

And the Bayes risk of this is

$$\begin{aligned}r(\hat{\lambda_i},\lambda_i) =&E^{\lambda_i|\mathbf{y},\gamma,\mu}\{\hat{\lambda_i}/\lambda_i-\log(\hat{\lambda_i}/\lambda_i)-1\}\\ =&\hat{\lambda_i}\cdot E^{\lambda_i|\mathbf{y},\gamma,\mu}(\lambda_i^{-1})-\log\hat{\lambda}_i+E^{\lambda_i|\mathbf{y},\gamma,\mu}(\log\lambda_i)-1.\end{aligned}$$

For the HB estimator for λ, $Y_1,\ldots,Y_p$ are marginally independent with density,

$$\begin{aligned}m(y_i|\gamma,\mu) = &\int[Y_i|\lambda_i][\lambda_i|\gamma,\mu]d\lambda_i\\ =&\frac{\Gamma(\alpha+n_i)y_i^{n_i-1}}{\Gamma(n_i)\Gamma(\alpha)\beta^\alpha(y_i+\beta^{-1})^{\alpha+n_i}}\\ =&\frac{\Gamma(1/\gamma+n_i)y_i^{n_i-1}\mu^{1/\gamma}}{\Gamma(n_i)\Gamma(1/\gamma)\gamma^{1/\gamma}(y_i+\mu/\gamma)^{1/\gamma+n_i}} \qquad (8)\end{aligned}$$

The HB estimator of λ is given by $\delta^{HB}(\mathbf{y})$ defined componentwise as

$$\begin{aligned}\delta_i^{HB}(\mathbf{y}) =& E(\lambda_i|\mathbf{y},\mu)\\ =& \tilde{\gamma}_i\bar{x}_i + (1-\tilde{\gamma}_i)\mu \end{aligned} \tag{9}$$

where

$$\tilde{\gamma}_i = E[\frac{n_i}{1/\gamma + n_i}|\mathbf{y},\mu].$$

The expectation is taken over the posterior distribution of γ given by

$$\begin{aligned}\pi(\gamma|\mathbf{y},\mu) &\propto \prod_{i=1}^{p} m(y_i|\gamma,\mu)\cdot\pi_2(\gamma)\\ &= \prod_{i=1}^{p} \frac{\Gamma(1/\gamma+n_i)y_i^{n_i-1}}{\Gamma(n_i)\Gamma(1/\gamma)\beta^{1/\gamma}(y_i+\beta^{-1})^{1/\gamma+n_i}}\cdot\pi_2(\gamma).\end{aligned}$$

Now, we consider the EB estimator for λ. The necessary method for EB approach is to maximize $m(\mathbf{y}|\gamma,\mu)$.

$$\begin{aligned}L(\gamma) &= \log m(\mathbf{y}|\gamma,\mu)\\ &= \log\prod_{i=1}^{p} m(y_i|\gamma,\mu)\\ &= \sum_{i=1}^{p}\log m(y_i|\gamma,\mu)\\ &= \sum_{i=1}^{p}[\log\Gamma(\gamma^{-1}+1) + (n_i-1)\log y_i + \gamma^{-1}\log\mu\\ &\quad -\log\Gamma(n_i) - \log\Gamma(\gamma^{-1}) - \gamma^{-1}\log\gamma\\ &\quad -(\gamma^{-1}+n_i)\log(y_i+\frac{\mu}{\gamma})].\end{aligned}$$

Therefore

$$\frac{\partial L}{\partial \gamma} = -\frac{1}{\gamma^2}[\sum_{i=1}^{p}\frac{\Gamma'(\gamma^{-1}+n_i)}{\Gamma(\gamma^{-1}+n_i)} + p\log(\mu/\gamma) - p\frac{\Gamma'(\gamma^{-1})}{\Gamma(\gamma^{-1})}$$

$$+ p - \sum_{i=1}^{p}\log(y_i+\mu/\gamma) - \sum_{i=1}^{p}\frac{\mu/\gamma}{y_i+\mu/\gamma}].$$

If L is maximized at neither 0 nor 1, the maximum can be found by setting above expression equal to 0. By using the solution of the equation $\hat{\gamma}$, we can find an EB estimator, given δ^{EB}, defined componentwise as

$$\delta_i^{EB}(\mathbf{y}) = \frac{n_i}{1/g\hat{m}+n_i}\bar{x}_i + [1 - \frac{n_i}{1/\hat{\gamma}+n_i}]\mu$$

$$= \frac{\hat{\gamma}n_i}{1+\hat{\gamma}n_i}\bar{x}_i + \frac{1}{1+\hat{\gamma}+n_i}\mu.$$

This approach ignores the error with respect to the estimation of the hyper-parameters. But the HB approach models the uncertainty of the hyperparameters by the second stage prior. This is the main advantage of the HB approach over the EB approach.

3. Comparison of HB and EB Estimators

In this section, we compare the HB and EB estimators that are found in previous section. First, we consider the shrinkage. The i-th component shrinkage of δ^{HB} and δ^{EB} away from the MVUE and invariant estimator $\bar{x}_i$

toward the prior mean μ are following respectively

$$\frac{\delta_i^{HB} - \bar{x}_i}{\mu - \bar{x}_i} = \frac{(1-\tilde{\gamma}_i)(-x_i+\mu)}{\mu - \bar{x}_i}$$
$$= 1 - \tilde{\gamma}_i$$
$$\frac{\delta_i^{EB} - \bar{x}_i}{\mu - \bar{x}_i} = [\frac{-1}{1+\hat{\gamma}n_i}\bar{x}_i + \frac{1}{1+\hat{\gamma}n_i}\mu]/(\mu - \bar{x}_i)$$
$$= 1/(1+\hat{\gamma}n_i)$$

where $i = 1, 2, \ldots, p$. Hence, both δ^{HB} and δ^{EB} shrink the UMVUE $\bar{\mathbf{X}}$ toward the prior mean μ.

As further study, Simulation study is needed for comparison of δ^{HB}, δ^{EB} and the UMVUE $\bar{\mathbf{X}}$. Moreover, we need to compare them with other estimator for θ or λ.

References

[1] Albert, J. H. (1981). Simultaneous estimation of Poisson means. *J. Multivariate Anal.* 11, 400-417.

[2] Brown, (1966) On the admissibility of invariant estimators of one or more location parameters. *Ann. Math. Statist.* 37, 1087-1136.

[3] Berger, J. (1980).Improving on inadmissible estimators in continuous exponential families with application to estimation of gamma scale parameter. *Ann. Statist.* 8, 545-571.

[4] Das Gupta, A. (1986). Simultaneous estimation in the multiparameter gamma distribution under weighted quadratic losses. *Ann. Statist.* 14, 206-219.

[5] Dey, D.K., M. Ghosh and C. Srinivasan (1987). Simultaneous estimation of parameters under entropy loss. *J. Statistical Planning and Inference.* 15, 347-363.

[6] Ghosh, M. and A. Parsian (1980). Admissible and minimax multiparameter estimation in exponential families. *J. Multivariate Anal.* 10, 551-564.

[7] Ghosh, M. and R. Auer (1983). Shrinkage estimates of multiple gamma paramters. Technical Report No. 202, Dept. of Statistics, Univ. of Florida, Gainesville.

[8] Ghosh, M., Hwang, J.T. and K.W. Tsui (1984). Construction of improved estimators in multiparameter estimation for continuous exponential families. *J. Multivariate Anal.* 14, 212-220.

[9] Ghosh, M. and D. K. Dey (1984) Trimmed estimates in simultaneous estimation of parameters in exponential families. *J. Multivariate Anal.* 15, 183-200.

[10] James, W. and C. Stein (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* Vol. 1, 361-379. University of California Press.

[11] Pepple, P.A. (1988). Simultaneous estimation of exponential means using Bayesian Methods. *Commun. Statist.-Theory Meth.* 17(2), 357-374.

[12] Stein, C. (1959). Admissibility of Pitman's estimator of a single location parameter. *Ann. Math. Statist.* 30, 970-979.