Calculation of Exact Estimators by Integration Over the Surface of an n-Dimensional Sphere

Anthony J. Webster*

May 6, 2013

Abstract

This paper reconsiders the problem of calculating the expected set of probabilities $\langle p_i \rangle$, given the observed set of items $\{m_i\}$, that are distributed among n bins with an (unknown) set of probabilities $\{p_i\}$ for being placed in the ith bin. The problem is often formulated using Bayes theorem and the multinomial distribution, along with a constant prior for the values of the p_i, leading to a Dirichlet distribution for the $\{p_i\}$. The moments of the p_i can then be calculated exactly. Here a new approach is suggested for the calculation of the moments, that uses a change of variables that reduces the problem to an integration over a portion of the surface of an n-dimensional sphere. This greatly simplifies the calculation by allowing a straightforward integration over $(n-1)$ independent variables, with the constraints on the set of p_i being automatically satisfied. For the Dirichlet and similar distributions the problem simplifies even further, with the resulting integrals subsequently

*email: dr.anthony.webster@gmail.com

factorising, allowing their easy evaluation in terms of Beta functions. A proof by induction confirms existing calculations for the moments. The advantage of the approach presented here is that the methods and results apply with minimum or no modifications to numerical calculations that involve more complicated distributions or non-constant prior distributions, for which cases the numerical calculations will be greatly simplified.

1 Introduction

Many problems involve placing N objects into n bins, with probabilities p_i for the object being placed into the ith bin. Given the values of the set of $\{p_i\}$, the probability density $P(m_1, m_2, ..., m_n|p_1, p_2, ..., p_n)$ for the distribution of the set of $\{m_i\}$ objects can be calculated, and is well-know as the multinomial distribution,

$$P(m_1, m_2, ..., m_n|p_1, p_2, ..., p_n) = \frac{N!}{m_1!m_2!...m_n!}\Pi_{i=1}^{n} p_i^{m_i} \tag{1}$$

with the constraint that $\sum_{i=1}^{n} p_i = 1$ and $\sum_{i=1}^{n} m_i = N$. Bayes theorem, $P(A|B)P(B) = P(B|A)P(A)$ requires,

$$P(p_1, p_2, ..., p_n|m_1, m_2, ..., m_n)P(m_1, m_2, ..., m_n) = P(m_1, m_2, ..., m_n|p_1, p_2, ..., p_n)P(p_1, p_2, ..., p_n) \tag{2}$$

that in principle allows us to calculate $P(p_1, p_2, ..., p_n|m_1, m_2, ...m_n)$, the probability of the set of probabilities $\{p_i\}$ with $i = 1$ to $i = n$, given the observed set of $\{m_i\}$. Often in such problems, $P(p_1, p_2, ..., p_n)$ is taken to be constant, and $P(m_1, m_2, ..., m_n)$ is chosen to ensure that $P(p_1, p_2, ..., p_n)$ is correctly normalised [1]. Applying this approach to the multinomial distribution, leads to a Dirichlet distribution, for which exactly calculated moments can be obtained. A recent approach to this problem by Friedman [2], relied on an identity discovered by Gauss that involves the integral representation of the hypergeometric distribution. The same is true of a recent exact calculation that corrects conjectured but widely used mark-recapture estimates [3], this and the coincidental timing of its revision on arXiv are what brought this problem to the author's attention.

Here an alternative method of calculation is considered. I suggest a change of variables that elegantly leads to a simple calculation for the moments of the $\{p_i\}$, and confirms existing results. The advantage of the method is that it can be applied very generally, and allows comparatively straightforward numerical integrations for the most general situations when analytical solutions may not be possible. The crux of the problem is the integration of a function over all possible values of p_i between 0 and 1, subject to the constraint of $\sum_{i=1}^{n} p_i = 1$. This appears in many situations, the specific case considered here is the product $\Pi_{i=1}^{n} p_i^{m_i}$ that arises in the Binomial, Multinomial, and Dirichlet distributions for example.

2 The Calculation

Consider the integration of the product $\Pi_{i=1}^{n} p_i^{m_i}$, over all sets of values of the p_i, subject to the constraints of $0 \leq p_i \leq 1$ for all i, and $\sum_{i=1}^{n} p_i = 1$. In Casella and Berger [4], the moments are obtained by a delightful trick (page 181), that simplifies the problem to integration over a binomial distribution. In Friedman [2] the integral is accomplished by a nested set of integrals, each of which depends on the calculation of the integrals within it, with for $n = 3$ for example,

$$I_3 = \int_{p_1=0}^{1} dp_1 \int_{p_2=0}^{1-p_1} dp_2 p_1^{m_1} p_2^{m_2} \left(1 - p_1 - p_2\right)^{m_3} \tag{3}$$

where $\sum_{i=1}^{3} p_i = 1$ has been used to write $p_3 = 1 - p_1 - p_2$. Here I start in a similar way, writing,

$$\Pi_{i=1}^{n} p_i^{m_i} = \left(1 - \sum_{i=1}^{n-1} p_i\right)^{m_n} \Pi_{i=1}^{n-1} p_i^{m_i} \tag{4}$$

that for $n = 3$ is $p_1^{m_1} p_2^{m_2} (1 - p_1 - p_2)^{m_3}$. Eq. 4 recognises that the constraint of $\sum_{i=1}^{n} p_i = 1$ leads to $(n-1)$ free parameters, or 2 free parameters for $n = 3$. For a radius of $r = 1$ the n-dimensional polar co-ordinates are:

$$\begin{aligned}
x_1(n) &= \cos\theta_1 \\
x_2(n) &= \sin\theta_1 \cos\theta_2 \\
x_3(n) &= \sin\theta_1 \sin\theta_2 \cos\theta_3 \\
... &\quad ... \\
x_{n-1}(n) &= \sin\theta_1 \sin\theta_2 ... \sin\theta_{n-2} \cos\theta_{n-1} \\
x_n(n) &= \sin\theta_1 \sin\theta_2 ... \sin\theta_{n-2} \sin\theta_{n-1}
\end{aligned} \tag{5}$$

Notice that $x_i(n)$ and $x_i(n)^2$ will vary continuously between 0 and 1 as the set of θ_i are varied continuously between 0 and $\pi/2$. Also notice that $\sum_{i=1}^{n} x_i(n)^2 = 1$, and consequently that $x_n(n)^2 = 1 - \sum_{i=1}^{n-1} x_i(n)^2$. Therefore the substitutions of $p_1 = x_1(n)^2$, $p_2 = x_2(n)^2$, ... , $p_{n-1} = x_{n-1}(n)^2$, will ensure that $\sum_{i=1}^{n} p_i = 1$, and integrals over θ_i from $\theta_i = 0$ to $\pi/2$ will allow p_i to vary continuously over all values between 0 and 1.

Note that the constraint of $\sum_{i=1}^{n} p_i = 1$ leads to $(n-1)$ free parameters, that after the change of variables, correspond to the set of θ_i with $i = 1$ to $(n-1)$. Also note that although we are using polar co-ordinates in n dimensions, because we have set $r = 1$, there are only $(n-1)$ free parameters.

The Jacobian of the co-ordinate transformation is $J = |\partial x_i(n)^2 / \partial \theta_j|$. Notice from Eq. 5

that $\partial x_i(n)^2/\partial\theta_j = 0$ for $j > i$. Consequently the determinant has zeros above the diagonal, and will evaluate easily to give $J = \Pi_{i=1}^{n-1} |\partial x_i(n)^2/\partial\theta_i|$.

Before proceeding to the general case, consider again the case with $n = 3$, for which case,

$$\begin{aligned} x_1(3) &= \cos\theta_1 \\ x_2(3) &= \sin\theta_1\cos\theta_2 \\ x_3(3) &= \sin\theta_1\sin\theta_2 \end{aligned} \tag{6}$$

The product $\left(1 - \sum_{i=1}^{n-1} p_i\right)^{m_n} \Pi_{i=1}^{n-1} p_i^{m_i}$ becomes, after the change of variables,

$$\begin{aligned} (1-p_1-p_2)^{m_3}\, p_1^{m_1} p_2^{m_2} &= \left(\sin^2\theta_1\sin^2\theta_2\right)^{m_3}\left(\cos^2\theta_1\right)^{m_1}\left(\sin^2\theta_1\cos^2\theta_2\right)^{m_2} \\ &= \left(\cos^{2m_1}\theta_1\sin^{2(m_2+m_3)}\theta_1\right)\left(\cos^{2m_2}\theta_2\sin^{2m_3}\theta_2\right) \end{aligned} \tag{7}$$

The Jacobian is,

$$\begin{aligned} J &= \begin{vmatrix} -2\cos\theta_1\sin\theta_1 & 0 \\ 2\sin\theta_1\cos\theta_1\cos^2\theta_2 & -2\sin^2\theta_1\sin\theta_2\cos\theta_2 \end{vmatrix} \\ &= \left(2\cos\theta_1\sin^3\theta_1\right)\left(2\sin\theta_2\cos\theta_2\right) \end{aligned} \tag{8}$$

Therefore using Eqs. 7 and 8 the integral in Eq 3 can be equivalently calculated from,

$$I_3 = \int_0^{\pi/2} d\theta_1 \int_0^{\pi/2} d\theta_2 \left(\cos^{2m_1}\theta_1\sin^{2(m_2+m_3)}\theta_1\right)\left(\cos^{2m_2}\theta_2\sin^{2m_3}\theta_2\right)\left(2\cos\theta_1\sin^3\theta_1\right)\left(2\sin\theta_2\cos\theta_2\right) \tag{9}$$

This integral factorises into,

$$I_3 = \left(2\int_0^{\pi/2} d\theta_1 \cos^{2(m_1+1)-1}\theta_1 \sin^{2(m_2+m_3+2)-1}\theta_1\right)\left(2\int_0^{\pi/2} d\theta_2 \cos^{2(m_2+1)-1}\theta_2 \sin^{2(m_3+1)-1}\theta_2\right) \tag{10}$$

the above Eq. 10 will be used as a starting point for a proof by induction for the general case later.

Many readers will immediately recognise the integrals as Beta functions, and it is well known that,

$$2\int_0^{\pi/2} d\theta \cos^{2m-1}\theta \sin^{2n-1}\theta = \mathsf{B}(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} \tag{11}$$

Consequently I_3 is easily evaluated as,

$$I_3 = \frac{\Gamma(m_1+1)\Gamma(m_2+m_3+2)}{\Gamma(m_1+m_2+m_3+3)}\frac{\Gamma(m_2+1)\Gamma(m_3+1)}{\Gamma(m_2+m_3+2)} \tag{12}$$

Cancelling terms and writing in terms of factorials this gives,

$$I_3 = \frac{m_1!m_2!m_3!}{(m_1+m_2+m_3+2)!} \tag{13}$$

For non-integer m_i Eq. 11 must be left written in terms of Gamma functions.

If we now wish to calculate $\langle p_1 \rangle$ for example, we simply need to evaluate $I_3(m_1{+}1, m_2, m_3)/I_3(m_1, m_2, m_3) = (m_1+1)/(m_1+m_2+m_3+3) = (m_1+1)/(N+3)$ with $N = m_1+m_2+m_3$, as found by Friedman. Other moments are easily calculated in a similar way.

For the general case, consider the formulae,

$$I_n = \int_0^{\pi/2} d\theta_1 \int_0^{\pi/2} d\theta_2 ... \int_0^{\pi/2} d\theta_{n-1} \Pi_{j=1}^{n-1} K_j(n) \tag{14}$$

$$K_j(n) = 2\cos^{2(m_j+1)-1}(\theta_j)\sin^{2\sum_{l=j+1}^{n}(1+m_l)-1}(\theta_j) \tag{15}$$

where I note that $\sum_{l=j+1}^{n}(1+m_l) = (n-j) + \sum_{l=j+1}^{n} m_l$, and the dependency on n of $K_j(n)$ is through the upper limit in the sum. Note that Eqs. 14 and 15 are true for $n=3$, as can be seen by comparison with Eq. 10. I will assume this is true for $n=k$ then show that this implies it is true for $n=k+1$, and consequently for all $k \geq 3$ by induction.

Firstly consider the integral with $n = k$. For $n = k$ the change of variables is,

$$\begin{aligned}
p_1 &= x_1(n)^2 &&= \cos^2\theta_1 \\
p_2 &= x_2(n)^2 &&= \sin^2\theta_1\cos^2\theta_2 \\
p_3 &= x_3(n)^2 &&= \sin^2\theta_1\sin^2\theta_2\cos^2\theta_3 \\
\ldots & &&\ldots \\
p_{k-1} &= x_{k-1}(n)^2 &&= \sin^2\theta_1\sin^2\theta_2\ldots\sin^2\theta_{k-2}\cos^2\theta_{k-1} \\
p_k &= x_k(n)^2 &&= \sin^2\theta_1\sin^2\theta_2\ldots\sin^2\theta_{k-2}\sin^2\theta_{k-1}
\end{aligned} \tag{16}$$

and the integrand is $\Pi_{i=1}^{k} p_i^{m_i}$, with a Jacobian that as noted previously, simplifies to $J = \Pi_{i=1}^{k-1}\left|\partial(x_i(k)^2)/\partial\theta_i\right|$. This gives the integral I_k as,

$$I_k = \int_0^{\pi/2} d\theta_1 \ldots \int_0^{\pi/2} d\theta_{k-1}\Pi_{i=1}^{k} x_i(k)^{2m_i}\Pi_{j=1}^{k-1}\left|\partial x_j(k)^2/\partial\theta_j\right| \tag{17}$$

Now consider $n = k+1$, for which the change of variables is,

$$\begin{aligned}
p_1 &= x_1(n)^2 &&= \cos^2\theta_1 \\
p_2 &= x_2(n)^2 &&= \sin^2\theta_1\cos^2\theta_2 \\
p_3 &= x_3(n)^2 &&= \sin^2\theta_1\sin^2\theta_2\cos^2\theta_3 \\
\ldots & &&\ldots \\
p_{k-1} &= x_{k-1}(n)^2 &&= \sin^2\theta_1\sin^2\theta_2\ldots\sin^2\theta_{k-2}\cos^2\theta_{k-1} \\
p_k &= x_k(n)^2 &&= \sin^2\theta_1\sin^2\theta_2\ldots\sin^2\theta_{k-2}\sin^2\theta_{k-1}\cos^2\theta_k \\
p_{k+1} &= x_{k+1}(n)^2 &&= \sin^2\theta_1\sin^2\theta_2\ldots\sin^2\theta_{k-2}\sin^2\theta_{k-1}\sin^2\theta_k
\end{aligned} \tag{18}$$

and the integral I_{k+1} is,

$$I_{k+1} = \int_0^{\pi/2} d\theta_1 \ldots \int_0^{\pi/2} d\theta_k\Pi_{i=1}^{k+1} x_i(k+1)^{2m_i}\Pi_{j=1}^{k}\left|\partial x_j(k+1)^2/\partial\theta_j\right| \tag{19}$$

Now notice that for $i = 1$ to $i = (k-1)$, $x_i(k) = x_i(k+1)$. For $i = k$, $x_k(k+1) = x_k(k)\cos^2\theta_k$.

Therefore,

$$\begin{aligned}\Pi_{i=1}^{k+1} x_i(k+1)^{2m_i} &= \Pi_{i=1}^{k} x_i(k)^{2m_i} \cos^{2m_k}(\theta_k) x_{k+1}(k+1) \\ &= \Pi_{i=1}^{k} x_i(k)^{2m_i} \cos^{2m_k}(\theta_k) \sin^{2m_{k+1}}(\theta_1) \sin^{2m_{k+1}}(\theta_2) ... \sin^{2m_{k+1}}(\theta_k)\end{aligned} \tag{20}$$

Similarly the Jacobian can be written as,

$$\begin{aligned}J &= \Pi_{i=1}^{k} \left|\tfrac{\partial}{\partial\theta_i}\left(x_i(k+1)^2\right)\right| \\ &= \left|\tfrac{\partial}{\partial\theta_k}\left(x_k(k+1)^2\right)\right| \Pi_{i=1}^{k-1} \left|\tfrac{\partial}{\partial\theta_i}\left(x_i(k)^2\right)\right| \\ &= -2\sin^2(\theta_1)\sin^2(\theta_2)...\sin^2(\theta_{k-1})\sin(\theta_k)\cos(\theta_k)\Pi_{i=1}^{k-1}\left|\tfrac{\partial}{\partial\theta_i}\left(x_i(k)^2\right)\right|\end{aligned} \tag{21}$$

Therefore we have,

$$\begin{aligned}I_{k+1} = \textstyle\int_0^{\pi/2} d\theta_1 ... \int_0^{\pi/2} d\theta_{k-1} \int_0^{\pi/2} d\theta_k \Pi_{i=1}^{k} x_i(k)^2 \Pi_{i=1}^{k-1}\left|\tfrac{\partial x_i(k)^2}{\partial\theta_i}\right| \\ \sin^{2(m_{k+1}+1)}(\theta_1)...\sin^{2(m_{k+1}+1)}(\theta_{k-1}) 2\cos^{2(m_{k+1}+1)-1}(\theta_k)\sin^{2(m_{k+1}+1)-1}(\theta_k)\end{aligned} \tag{22}$$

Comparing Eq. 17 with the assumption of Eq. 14, we find,

$$\Pi_{i=1}^{k} x_i(k)^{2m_i} \Pi_{i=1}^{k-1}\left|\frac{\partial x_i(k)^2}{\partial\theta_i}\right| = \Pi_{i=1}^{k-1} K_j(k) \tag{23}$$

with $K_j(k)$ given by Eq. 15. Under this assumption the integrand of Eq. 22 can be written as,

$$\left[2\cos^{2(m_{k+1}+1)-1}(\theta_k)\sin^{2(m_{k+1}+1)-1}(\theta_k)\right]\Pi_{i=1}^{k-1}\left[K_j(k)\sin^{2(m_{k+1}+1)}(\theta_j)\right] \tag{24}$$

Note that,

$$\begin{aligned}K_j(k)\sin^{2(m_{k+1}+1)}(\theta_j) &= 2\cos^{2(m_j+1)-1}(\theta_j)\sin^{2\sum_{l=j+1}^{k+1}(1+m_l)-1}(\theta_j) \\ &= K_j(k+1) \text{ for } 1 \le j \le (k-1)\end{aligned} \tag{25}$$

The extra factor in Eq. 24 is,

$$2\cos^{2(m_{k+1}+1)-1}(\theta_k)\sin^{2(m_{k+1}+1)-1}(\theta_k) = K_k(k+1) \tag{26}$$

Therefore we have,

$$I_{k+1}=\int_0^{\pi/2}d\theta_1...\int_0^{\pi/2}d\theta_k\Pi_{i=1}^{k}K_i(k+1) \tag{27}$$

which is just Eq. 14 with $n=(k+1)$, and $K_i(k+1)$ as given by Eq. 15. Since we've shown Eq. 27 to be true for $n=3$ and that its truth for $n=k$ implies it to be true for $n=(k+1)$, then by induction Eqs. 14 and 15 are true for all $n\geq 3$.

Eq. 27 is easy to evaluate. Because θ_i only appears in $K_i(k+1)$, the integral factors into,

$$I_{k+1}=\Pi_{i=1}^{k}\int_0^{\pi/2}d\theta_i K_i(k+1) \tag{28}$$

Noting Eq. 15 for $K_i(k+1)$, each of the integrals can be recognised as a Beta function, with,

$$\begin{aligned}\int_0^{\pi/2}d\theta_i K_i(k+1) &= 2\int_0^{\pi/2}\cos^{2(m_i+1)-1}(\theta_i)\sin^{2\sum_{l=i+1}^{k+1}(1+m_l)-1}(\theta_i)\\ &= \frac{\Gamma(m_i+1)\Gamma\left(\sum_{l=i+1}^{k+1}(1+m_l)\right)}{\Gamma\left(\sum_{l=j}^{k+1}(1+m_l)\right)}\end{aligned} \tag{29}$$

where in the denominator of the last line we used $m_i+1+\sum_{l=i+1}^{k+1}(1+m_l)=\sum_{l=i}^{k+1}(1+m_l)$. To obtain an explicit value for the integral, now we simply need to multiply out the terms, with,

$$\begin{aligned}I_{k+1}=&\frac{\Gamma(m_1+1)\Gamma\left(\sum_{l=2}^{k+1}(1+m_l)\right)}{\Gamma\left(\sum_{l=1}^{k+1}(1+m_l)\right)}\times\frac{\Gamma(m_2+1)\Gamma\left(\sum_{l=3}^{k+1}(1+m_l)\right)}{\Gamma\left(\sum_{l=2}^{k+1}(1+m_l)\right)}\times...\\ &...\times\frac{\Gamma(m_{k-1}+1)\Gamma(m_k+m_{k+1}+2)}{\Gamma(m_{k-1}+m_k+m_{k+1}+3)}\times\frac{\Gamma(m_k+1)\Gamma(m_{k+1}+1)}{\Gamma(m_k+m_{k+1}+2)}\end{aligned} \tag{30}$$

Cancelling successive terms, leaves,

$$I_{k+1}=\frac{\Gamma(m_1+1)\Gamma(m_2+1)...\Gamma(m_k+1)\Gamma(m_{k+1}+1)}{\Gamma\left(\sum_{l=1}^{k+1}(1+m_l)\right)} \tag{31}$$

which when written in terms of factorials and $N=\sum_{l=1}^{k+1}m_l$, gives,

$$I_{k+1}=\frac{m_1!m_2!...m_{k+1}!}{(N+k)!} \tag{32}$$

For non-integral values of m_i the Eq. 32 must be remain expressed in terms of Gamma functions. Note that the above expression (32) is for $n = k + 1$, and usually we will evaluate it with $n = k$, for which case $I_k = m_1!m_2!...m_k!/(N + k - 1)!$.

To obtain the qth moment of p_i one simply needs to substitute $(m_i + q)$ for m_i in I_k, and calculate the ratio of $I_k(m_i + q)/I_k(m_i)$, whose meaning is hopefully clear. For example, $\langle p_i \rangle$ is given by,

$$\langle p_i \rangle = \frac{m_1!m_2!...(m_i + 1)!...m_k!}{(N + k)!} \frac{(N + k - 1)!}{m_1!m_2!...m_k!} = \frac{m_i + 1}{N + k} \tag{33}$$

where the notation $\langle p_i \rangle$ is used to denote the moment of p_i when there are k "bins". Similarly,

$$\langle p_i^2 \rangle = \frac{m_1!m_2!...(m_i + 2)!...m_k!}{(N + k + 1)!} \frac{(N + k - 1)!}{m_1!m_2!...m_k!} = \frac{(m_i + 2)(m_i + 1)}{(N + k + 1)(N + k)} \tag{34}$$

Giving the standard deviation as,

$$\langle p_i^2 \rangle - \langle p_i \rangle^2 = \frac{(m_i + 1)(N + k - m_i - 1)}{(N + k)^2(N + k + 1)} \tag{35}$$

These results are in agreement with those of Friedman. Higher order moments are also easily calculated. The difference of the skewness from zero for example, can give an indication of the extent to which noise in the data should be regarded as non-Gaussian. Note that because $\sum_{i=1}^n p_i = 1$, then,

$$\begin{aligned} 1 &= \int_D dp_1...dp_{k-1} \left(\sum_{i=1}^k p_i \right) P(p_1, ..., p_k | m_1, ..., m_k) \\ &= \sum_{i=1}^k \langle p_i \rangle \end{aligned} \tag{36}$$

where D is used as shorthand to indicate that the integral should be over the correct domain of integration subject to the constraint of $\sum_{i=1}^n p_i = 1$. Eq. 36 is correctly satisfied by Eq. 33.

3 Remarks

There are a variety of distributions in which the $\{p_i\}$ only appear in a factor of $\Pi_{i=1}^{n} p_i^{m_i}$, and the results here apply to those cases also. More generally the probability distribution or its prior could involve any function of $\{p_i\}$. For example, we might want to introduce a suitable prior into the problem so as to bias against "outliers", or towards a particular set of $\{p_i\}$. In these more general cases the change of variables to n-dimensional spherical polars will still allow a comparatively straightforward numerical integral. A numerical integral over the $\{p_i\}$ subject to $0 \leq p_i \leq 1$ and $\sum_{i=1}^{n} p_i = 1$, without the change of variables to spherical polars, is not so easy. For some combinations of priors and probability distributions the integral will remain factorisable after the change of variables. This might continue to be useful for other analytical calculations.

References

[1] E.T. Jaynes "Probability Theory The Logic of Science", Cambridge University Press, 2003.

[2] J.M. Friedman "Unbiased estimators for the parameters of the binomial and multinomial distributions", arXiv: 1302.5749v1.

[3] A.J. Webster and R. Kemp "Estimating Omissions from Searches" (arXiv: 1205.1150v2) The American Statistician, in press, (2013).

[4] G. Casella and R.L. Berger "Statistical Inference", second edition, (2002).