Lecture I: Some useful approximations

I.1. Notation

The notation we use will assume that we have a random vector $Y = (Y_1, \ldots, Y_n)$, where $Y_1, \ldots, Y_n$ are independent, identically distributed, from a density function $f(y)$ or $f(y; \theta)$ where $\theta \in \mathbb{R}^p$.

- moment generating function $M_Y(t) = E\ e^{t Y}$
- cumulant generating function $K_Y(t) = \log M_Y(t) = \kappa_1 t + \frac{1}{2} \kappa_2 t^2 + \frac{1}{6} \kappa_3 t^3 + \ldots$
- standardized sum $S_n = \sum Y_i; \quad S_n^* = (S_n - n\mu)/\sqrt{n}\sigma$
  where $\mu = \kappa_1 = EY_1, \quad \sigma^2 = \kappa_2 = \text{var} Y_1$
- standardized cumulants $\rho_3 = \kappa_3/\kappa_2^{3/2}; \quad \rho_4 = \kappa_4/\kappa_2^2$
- Hermite polynomials $H_r(x) = (-1)^r \phi^{(r)}(x)/\phi(x)$

*Kendall & Stuart, Vol. I; McCullagh for multivariate*
I.2 Edgeworth expansion

The Edgeworth expansion for the distribution of the sample mean can be expressed in the following equivalent ways:

1. 
\[ f_{s_n^*}(x) = \phi(x) \left\{ 1 + \frac{\rho_3}{6\sqrt{n}} H_3(x) + \frac{\rho_4}{24n} H_4(x) + \frac{\rho_5^2}{72n} H_6(x) \right\} + O(n^{-3/2}) \]

\[ F_{s_n^*}(x) = \Phi(x) - \phi(x) \left\{ \frac{\rho_3}{6\sqrt{n}} H_3(x) + \frac{\rho_4}{24n} H_4(x) + \frac{\rho_5^2}{72n} H_6(x) \right\} + O(n^{-3/2}) \]

2. 
\[ f_{s_n}(z) = \frac{1}{\sqrt{n\sigma}} \phi\left( \frac{z - n\mu}{\sqrt{n\sigma}} \right) \left\{ 1 + \frac{\rho_3}{6\sqrt{n}} H_3\left( \frac{z - n\mu}{\sqrt{n\sigma}} \right) + \frac{\rho_4}{24n} H_4\left( \frac{z - n\mu}{\sqrt{n\sigma}} \right) + \frac{\rho_5^2}{72n} H_6\left( \frac{z - n\mu}{\sqrt{n\sigma}} \right) \right\} + O(n^{-3/2}) \]

3. Let \( \phi(z; a, b^2) = (1/\sqrt{b}) \phi\left\{ (z - a)/b \right\} \), \( H_r(z; a, b^2) = (-1)^r \phi^{(r)}(z; a, b^2)/\phi(z; a, b^2) \), and let \( \kappa_r^n = \kappa_r(S_n) \) be the \( r \)th cumulant of \( S_n \).

\[ f_{s_n}(z) = \phi(z; n\mu, n\sigma^2) \left\{ 1 + \frac{\kappa_3^n}{6} H_3(z; n\mu, n\sigma^2) + \frac{\kappa_4^n}{24} H_4(z; n\mu, n\sigma^2) + \frac{\kappa_5^n}{72} H_6(z; n\mu, n\sigma^2) \right\} \]

4. The above are for scalar \( Y_i \); the multiparameter Edgeworth expansion is a generalization of 3. and takes the form

\[ f_{s_n}(z) = \phi(z; \mu, \Sigma) \left\{ 1 + \frac{1}{6} \kappa^{ijk} H_{ijk}(z; \mu, \Sigma) + \frac{1}{24} \kappa^{ijkld} H_{ijkl}(z; \mu, \Sigma) + \frac{1}{72} \kappa^{ijklm} [10] H_{ijkl}(z; \mu, \Sigma) H_{lmn}(z; \mu, \Sigma) \right\} \]

5. If \( T_n = t(S_n) \) is a sufficiently smooth function of \( S_n \) (or \( S_n^* \)), then a formal Edgeworth expansion for the density of \( T_n \) is valid.

_Feller; Barndorff-Nielsen & Cox 89; McCullagh; Kolassa; Skovgaard 81ab_
I.3 Saddlepoint expansion

The saddlepoint expansion for the density of $S_n$ is

$$
\hat{f}_{S_n}(s) = \frac{1}{\sqrt{2\pi}} \frac{1}{\{nK_Y'\hat{\phi}\}^{1/2}} \exp\{nK_Y(\hat{\phi}) - \hat{\phi}s\} \{1 + O(n^{-1})\}
$$

(1)

where $\hat{\phi}$ satisfies the equation $nK_Y'\hat{\phi} = s$

- the $O(n^{-1})$ term in (1) is $(3\hat{p}_4 - 5\hat{p}_2^2)/(24n)$ where $\hat{p}_j = p_j(\hat{\phi})$

- If the leading term of (1) is renormalized it approximates the density of $S_n$ with relative error $O(n^{-3/2})$

- a simple change of variables gives a saddlepoint expansion for $\tilde{S}_n = S_n/n$

- if $Y_i$ are $p$-dimensional vectors, then $M_Y(t) = E\exp(t^TY)$ and

$$
\hat{f}_{S_n}(s) = \frac{c}{\sqrt{2\pi}} \frac{1}{\{n|K_Y'\hat{\phi}|\}^{1/2}} \exp\{nK_Y(\hat{\phi}) - \hat{\phi}^T s\}
$$

(2)

where $K_Y'\hat{\phi}$ is a vector and $K_Y''(\phi)$ is a $p \times p$ matrix.

- If $f(y) = \exp\{\theta^Ty - k(\theta) - d(y)\}$, then $K(\phi) = k(\theta + \phi) - k(\theta), \hat{\phi} = \hat{\theta} - \theta$, and (1) becomes

$$
\hat{f}_{\hat{\phi}}(\hat{\theta}) = \frac{c}{\sqrt{2\pi}} |n\theta(\hat{\theta})|^{1/2} \exp\{\ell(\theta) - \ell(\hat{\theta})\}
$$

(3)

using the change of variable $nk'(\hat{\theta}) = s$. In (3) $\ell(\theta) = \ell(\theta; y_1, \ldots, y_n) = \ell(\theta; \hat{\theta}) + C = \theta^Ts - nk(\theta) = \theta^T nk'(\hat{\theta}) - nk(\theta)$.

BNC 89; McCullagh; Kolassa; Field & Ronchetti
I.4 Laplace approximation of integrals

1. \[
\int_a^b e^{-n\varphi(y)} \, dy = e^{-n\varphi(\bar{y})} \sqrt{\frac{2\pi}{n}} \left( g''(\bar{y}) \right)^{-1/2} \left\{ 1 + \frac{5\hat{\rho}_3^2 - 3\hat{\rho}_4}{24n} + O(n^{-2}) \right\}
\]

where \( g'(\bar{y}) = 0, \, g''(\bar{y}) > 0, \, \hat{\rho}_3 = g'''(\bar{y}) / \{g''(\bar{y})\}^{3/2}, \, \hat{\rho}_4 = g^{(4)}(\bar{y}) / \{g''(\bar{y})\}^2, \) and we assume here and in the following that the function \( g \) has a unique non-zero minimum in the interval \((a, b)\).

2. \[
\int h(y)e^{-n\varphi(y)} \, dy = h(\bar{y})e^{-n\varphi(\bar{y})} \sqrt{\frac{2\pi}{n}} \left( g''(\bar{y}) \right)^{-1/2} \left\{ 1 + \frac{5\hat{\rho}_3^2 - 3\hat{\rho}_4}{24n} + \frac{h''(\bar{y})}{2g''(\bar{y})} - \frac{\hat{\rho}_3 h'(\bar{y})}{h(\bar{y})} + O(n^{-2}) \right\}
\]

for which we need to assume that \( h(\bar{y}) \neq 0 \).

3. \[
\int e^{-n\{\varphi(y) - \frac{1}{2} \log h(y)\}} \, dy = \int e^{-n\varphi_n(y)} \, dy, \text{ say,}
\]

\[
e^{-n\varphi(y^*)} h(y^*) \sqrt{\frac{2\pi}{n}} \left( q''_n(y^*) \right)^{-1/2} \left\{ 1 + (5\rho_3^* - 3\rho_4^*)/(24n) + O(n^{-2}) \right\}
\]

where \( q'_n(y^*) = 0, \rho_j^* = q^{(j)}_n(y^*) / \{q''_n(y^*)\}^{j/2} \).

4. \[
\int_{Re} h(y)e^{-n\varphi(y)} \, dy = e^{-n\varphi(\bar{y})} h(\bar{y}) \left( \sqrt{\frac{2\pi}{n}} \right)^d |g''(\bar{y})|^{-1/2} \left\{ 1 + O(n^{-1}) \right\}
\]

\[
e^{-n\varphi(y^*)} h(y^*) \left( \sqrt{\frac{2\pi}{n}} \right)^d |g''(y^*)|^{-1/2} \left\{ 1 + O(n^{-1}) \right\}
\]

\textit{BNC 89; Tierney & Kadane 86}
1.5 Equivalence of $\chi^2$ approximations

Denote the density of the chi-squared distribution with $d$ degrees of freedom by $q_d(x)$. The following expansions are formally equivalent:

1. The density of $X_n$ is
   
   $$q_d(x)(1 - a/n) + q_{d+2}(x)an^{-1} + O(n^{-2}).$$

2. The density of $X_n$ is
   
   $$q_d(x)\{1 + a(x/d - 1)n^{-1}\} + O(n^{-2}).$$

3. The moment generating function of $X_n$ is
   
   $$(1 - 2t)^{-d/2}\{1 + 2at(1 - 2t)^{-1}n^{-1}\} + O(n^{-2}).$$

4. The density of $\{1 + 2a/(dn)\}X_n$ is $q_d(x)$, with error $O(n^{-2})$.

*BNC 89, Ch.4 (Exercises)*
Selected references

Contains proof of exactness of $p^*$ in transformation models.


Derivation of scalar parameter tail area approximation in the $r^*$ form.


Ch.2: First order theory for likelihood based quantities, delta method and Slutsky’s theorem. Ch.3: Laplace’s method. Ch.4: Edgeworth and saddlepoint


Ch.3: First order theory. Ch.6: $p^*$ and $r^*$ approximation. Ch.4.4 and 6.5: Derivation of Bartlett correction from $p^*$ formula. Ch.7.2: approximate ancillarity.


Ch.4: Sketch of asymptotic normality of posterior.


Careful derivation of the Bartlett correction for the posterior log-likelihood ratio statistic, and the use of an 'unsmoothing' argument due to Stein and to Ghosh, also outlined in Dawid (1991, JRSSB), to obtain the frequentist Bartlett correction result. Does not give an explicit form for the Bartlett factor, but this is available for the Bayesian case in DS93 and for the frequentist case in BNC94, Ch.6.

With F&McD84, generalizes Walker’s result on asymptotic normality of the posterior, to improper priors, and simplifies the proof.


Ch.9: First order theory for likelihood based quantities, with and without nuisance parameters. Bartlett correction.


Derivation of tail area approximation with nuisance parameters in transformation (location-scale) models.


Applies the tail area approximation of DFF90 to the Laplace approximation for the marginal posterior of a nonlinear function. The latter extends Tierney and Kadane (1986) and is derived in TKK89. Section 4 gives an application to conditional inference in exponential families.


Uses Bayesian tail area approximation to find matching priors.


Derivation of Bayesian Bartlett correction by brute force, starting from the Laplace approximation to the marginal posterior.

Ch. 16: Clear and careful proof of the Edgeworth expansion.


Special emphasis on density approximations for M-estimates.


Derivation of scalar parameter tail area approximation in the Lugannani and Rice form. Lots of numerical examples.


Scalar parameter tail area approximation and a discussion of the confidence distribution function, or significance function.


Derivation of tail area approximation for conditional inference in exponential families.


Edgeworth expansion for the posterior cdf, in detail.

Very good review of reference priors.


Provides a very detailed derivation of the Edgeworth and saddlepoint expansions for the density and distribution function of the sample mean.


The original and most impressive paper on Bartlett correction.


Ch. 6: Asymptotic consistency and normality of the m.l.e., using a 'stochastic Taylor series’ argument. Good discussion of conditions.


Review of $r^*$ in exponential families, with detailed numerical work.


A review of the $p^*$ formula, in exponential and transformation families, and the relationship to the saddlepoint approximation.


Brief review of matching priors in Section 4; lots of references.


Ch.4.4.4: First order theory for log-likelihood ratio statistics.


These papers give a careful outline of the use of the Edgeworth expansion for the sample mean to derive an Edgeworth expansion for a smooth function of a sample mean.


Beautifully written, includes a proof of the $p^*$ formula in Section 3.


The source for everything you ever want to know about cumulants and moments for scalar random variables, in Ch. 3; Edgeworth expansion in detail in Ch. 6.


Detailed proof of asymptotic normality of posterior.


Derivation of the Jeffreys' prior as a 'matching' prior, in the sense of giving posterior probability intervals with the correct frequentist coverage, to $O(n^{-1})$, which seems to be the best that can be done using a Bayesian approach. Scalar parameter case only. Expression given for the $O(n^{-1})$ term.