3. Laplace's method

a. Suppose that data arise as n independent and approximately identically distributed observations

i. Treat $\ell(\theta) = \log(L(\theta))$ as $n\ell^\circ(\theta)$

- ℓ° depends on n through the data
- Heuristically ignore this dependence on n .

b. We want $I_A = \int_A \exp(n\ell^\circ(\theta))\varpi(\theta)\,d\theta$

c. Let $\hat\theta$ be the MLE.

d. Do Taylor series approximation for ℓ° and $\varpi(\theta)$ separately.

i. $I_A \approx \exp(n\ell^\circ(\hat\theta))\int_A \exp(n{\ell^\circ}''(\theta)(\theta-\hat\theta)^2/2)(1 + n{\ell^\circ}'''(\hat\theta)(\theta-\hat\theta)^3/3)(\varpi(\hat\theta)+\varpi'(\hat\theta)(\theta-\hat\theta))\,d\theta$

e. Reparameterize to remove sample size from exponent.

i. Let $\vartheta = \sqrt{-{\ell^\circ}''}(\theta-\hat\theta)/\sqrt{n}$, $c_1 = {\ell^\circ}'''(\hat\theta)(-{\ell^\circ}''(\hat\theta))^{-3/2}$, $c_2 = \varpi'(\hat\theta)(-{\ell^\circ}''(\hat\theta))^{-1/2}/\varpi(\hat\theta)$ and $A' = \{(\theta-\hat\theta)/\sqrt{-{\ell^\circ}''(\hat\theta)}|\theta\in A\}$..

ii. Then $I_A \approx \frac{\sqrt{n}\exp(n\ell^\circ(\hat\theta))\varpi(\hat\theta)}{\sqrt{{\ell^\circ}''(\hat\theta)}}\int_{A'}\exp(-\vartheta^2/2)(1 + c_1\vartheta^3/(6\sqrt{n}))(1+c_2\vartheta/\sqrt{n}))\,d\vartheta$.

iii. Denominator is $I_{(-\infty,\infty)} = \sqrt{n}\exp(n\ell^\circ(\hat\theta))\varpi(\hat\theta)\sqrt{-{\ell^\circ}''(\hat\theta)/(2\pi}$

iv. Then P $[\theta \le c|\text{data}]$ is approximately

$$\Phi\left(\sigma(c-\hat{\theta})\right) - \frac{\phi(\sqrt{n}\sigma(c-\hat{\theta}))}{\sqrt{n}}\left[c_1\left(n\sigma^2(c-\hat{\theta})^2+2\right)/6-c_2\right]$$

v. So with sample size increasing, prior kept constant, posterior is approximately Gaussian,

vi. and expectation and standard deviation of approximating distribution are the same as in frequentist likelihood inference.

vii. Largest deviations from approximate normality are driven by asymmetry of prior and likelihood about MLE.

f. Extends to multiple dimensions and higher-order approximatons.

g. Example: Inference on Binomial Proportion, Beta Prior

i. Log likelihood $\ell(\pi) = X\log(\pi) + (n-X)\log(1-\pi)$.

ii. $\ell'(\pi) = X/\pi - (n-X)/(1-\pi)$.

iii. $\ell''(\pi) = -X/\pi^2 - (n-X)/(1-\pi)^2$.

iv. $\hat{\pi} = X/n$

v. $-\ell''(\hat{\pi}) = n/((X/n)(1-X/n))$.

vi. P $[\pi < c|X] = \Phi((c-X/n)\sqrt{n/((X/n)(1-X/n))})$.

4. Many integration methods scale poorly with θ dimension

a. Deterministic integration is replaced by simulation. R Code

A: 5.4-5.4.2

5. Bayesian Methods for Regression Models

   a. Use existing generalized linear model for likelihood.

   b. Put prior on model parameters.

      i. Regression parameters take values in $(-\infty, \infty)$.

      - Priors should put positive probability in all plausible regions of this line.
      - One might consider a prior flat in the same sense as the prior uniform on $(0, 1)$.
      - Any prior density that is constant over the real line cannot integrate to a finite quantity.
      - Such flat priors on the real line are examples of *improper priors*.

      ii. Other model parameters also need priors

      - Ex., dispersion parameter in error distribution for continuous response variables.

      iii. R Code

: