Bayesian STAT Test # 1 Review

- Prior: the ~~[illegible]~~ belief you have about the underlying distribution.

* we start with our prior probability dist. over some parameter, then we use data to update that distribution to become the posterior distribution ~~[illegible]~~ that expresses our new belief. *

- Yesterdays Posterior is todays prior
- ~~Posterior Mean = Prior Mean~~
- The 3 conjugate priors are:
  - Beta - Binomial
  - Gamma - Poisson
  - Normal - Normal
- Confidence Intervals v.s. Credible Intervals:
  - Confidence intervals are frequenist & Credible intervals are Bayesian.

"95% of similarly constructed intervals will contain the true mean"

"The probability that the true mean is contained within the given interval is 95%"

- Classical / frequenist approach: use the sample to make inference about the unknown params.
- Bayesian approach: treat unknown params as RV's having a specified distribution.
- Bayes Estimate = $\hat{\theta} = E(\theta | X) = \frac{a_1}{b_1}$ ← Gamma
- Gibbs requires conjugate priors but is accurate
- Metropolis - Hastings sampling is flexible but requires tuning.

Conjugate priors:

$\pi(P|X) = P(X|P)\,\pi(P)$

$\pi(P|X) =$ known dist. PDF x prior PDF

√ Normal likelihood:

$$f(x|\mu) = \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

√ Binomial ~~PDF~~ Likelihood:

$$f(x|\theta) = \binom{n}{k}\theta^k(1-\theta)^{n-k}$$

√ Poisson Likelihood:

$$f(x|\lambda) = \cancel{\frac{\lambda^x}{x!}} \quad \cancel{\exp(-N\lambda)\lambda^x} \quad \cancel{\frac{\lambda^x \exp(-\lambda)}{x!}}$$

$$\frac{\lambda^x \exp(-\lambda)}{x!}$$

Normal PDF:

√ $$\exp\left\{-\frac{1}{2\tau^2}(\mu-\delta)^2\right\} = \pi(\mu)$$

√ Beta PDF:

$$\pi(\theta) = \theta^{a-1}(1-\theta)^{b-1}$$

√ Gamma PDF:

$$\cancel{\exp}\ \pi(\lambda) = \exp(-B\lambda)\lambda^{a-1}$$

$\hat{\theta} = \frac{a}{a+b}$ Beta-Binomial: Posterior $\propto$ $\theta|Y \sim Beta(a+n, b+Y)$

$\hat{\theta} = a/b$ Gamma-Poisson: Poster $\propto$ $\lambda|Y \sim Gamma(a+\sum_{i=1}^{n} Y_i, b+\sum_{i=1}^{n} N_i)$

Normal-Normal: Posterior $\propto$ $\mu|Y \sim Normal(\quad)$

Chapter 3

Diagnostics.R

Poor Convergence signs:

- density fits poorly.
- Chains do not converge.
- Autocorrelation has not "dropped fast enough"
- Gelman "multivariate psrf" $R > 1.1$
- Geweke diag $|z| < 2$ indicates convergence
- ESS, the higher the better, should be over 1000.

How to improve convergence:

- Increase # of iterations
- Pick more informative priors
- Improve init values
- use a simpler ~~or more advanced algo~~ model
- use a more advanced algorithm

- If the # of params is ~~less than~~ Greater than the # of observations then the convergence will be super slow. Not all params are identifiable.

Chapter 4

Linear_Regression.R, ~~Bayes Factor.R~~ prediction_with_Linear_Regression.R

There are 3 different priors for the slopes in linear Regression: *03/08 for more detail*

- Uninformative Gaussian: $\beta_j \sim Normal(0, 1000)$
- Gaussian shrinkage: $\beta_j \sim Normal(0, \sigma_b^2)$ with $\sigma_b^2 \sim InverseGamma(0.1, 0.1)$
- Bayesian Lasso: $\beta_j \sim DE(0, \sigma_b^2)$ with $\sigma_b^2 \sim InverseGamma(0.1, 0.1)$

Are both Shrinkage Estimators. Are useful when we want a stabler estimate of β, β is needed when $p > n$. $p = ncol(X)$, $n = len(y)$

$$\beta = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} X_{ij} \beta_j\right)^2 + \lambda \sum_{j=1}^{p} \beta_j \right.$$

Chapter 5 | bayesFactors.R , model_selection.R

for AIC, BIC, DIC, & WAIC the smaller the better the model

DIC = Mean deviance + ~~Penalized deviance~~ Penalty = Penalized deviance

How well of a fit

balance of fit v.s. complexity

WAIC = WAIC

Bayes Factor. ~~over 10 is Stron~~ B_{12} with $B_{12} > 10$ is strong evidence for model 1. Used for hypothesis testing; provides evidence that the data supports one model over another. *A $B_{12} = X$ means that the data is X times more likely under the alternate hypothesis.

BF = P(Alternate) / P(null)

Posterior odds BF prior odds

$$\frac{\pi(M_1|Y)}{\pi(M_1|Y)} = \frac{f_1(Y)}{f_2(Y)} \cdot \frac{\pi(M_1)}{\pi(M_2)}$$