4.8 EXERCISES (Part 2)

BMED6420
Brani Vidakovic
Consult the class slides, hints, and cited literature for comprehensive solution of exercise problems.

Counts of Alpha Particles. Rutherford and Geiger provided counts of α-particles in their experiment. The counts, given in the table below, can be well modeled by Poisson distribution.

X	0	1	2	3	4	5	6	7	8	9	10	11	≥ 12
Frequency	57	203	383	525	532	408	273	139	45	27	10	4	2

(a) Find sample size n and sample mean $\bar{X}$. In calculations for $\bar{X}$ take ≥ 12 as 12.

(b) Elicit a gamma prior for λ with rate parameter $\beta = 10$ and shape parameter α selected in such a way that the prior mean is 7.

(c) Find the Bayes estimator of λ using the prior from (b). Is the problem conjugate? Use the fact that $\sum_{i=1}^{n} X_i \sim \mathcal{P}oi(n\lambda)$.

(d) Write a WinBUGS script that simulates the Bayes estimator for λ and compare its output with the analytic solution from (c).

Mosaic Virus. A single leaf is taken from each of 8 different tobacco plants. Each leaf is then divided in half, and given one of two preparations of *mosaic* virus. Researchers wanted to examine if there is a difference in the mean number of lesions from the two preparations. Here is the raw data:

Plant	Prep 1	Prep 2
1	38	29
2	40	35
3	26	31
4	33	31
5	21	14
6	27	37
7	41	22
8	36	25

Assume the normal distribution for the difference between the populations/samples.
Using WinBUGS/OpenBUGS find
(a) the 95% credible set for $\mu_1 - \mu_2$, and
(b) posterior probability of hypothesis $H_1 : \mu_1 - \mu_2 \geq 0$.
Use noninformative priors.

Hint. Since this is a paired two sample problem, a single model should placed on the difference.

FIGO. Despite the excellent prognosis of FIGO[1] stage I, type I endometrial cancers, a substantial number of patients experience recurrence and die from this disease.

[1] Federation Internationale de Gynecologie et d'Obstetrique

Zeimet et al (2013) conducted a retrospective multicenter cohort study to determine expression of L1CAM by immunohistochemistry in 1021 endometrial cancer specimens with the goal to predict clinical outcome.

Of 1021 included cancers, 17.7% were rated L1CAM-positive. Of these L1CAM-positive cancers, 51.4% recurred during follow-up compared with 2.9% L1CAM-negative cancers. Patients bearing L1CAM-positive cancers had poorer disease-free and overall survival.

It is stated that L1CAM has been the best-ever published prognostic factor in FIGO stage I, type I endometrial cancers and shows clear superiority over the standardly used multifactor risk score. L1CAM expression in type I cancers indicates the need for adjuvant treatment. This adhesion molecule might serve as a treatment target for the fully humanized anti-L1CAM antibody currently under development for clinical use.

	FIGO I/I Endometrial Cancer		
	Recurred	Did Not Recur	Total
L1CAM Positive			
L1CAM Negative			
Total			1021

(a) Using information supplied fill in the table (round the entries to the closest integer).

(b) The estimators of the population Sensitivity and Specificity are simple relative frequencies (ratios): True Positives (TP)/Recurred and True Negatives (TN)/Not Recurred. Consider now a Bayesian version of this problem. Using WinBUGS model TP and TN as Binomials, place the priors on population Sensitivity (p_1) and Specificity (p_2) and find their Bayesian estimators.

Explore the estimators for your favorite choice of priors on p_1 and p_2: Jeffreys', uniform (0,1), flat on logit, etc.

Zeimet, A. G., Reimer, D., Huszar, M., Winterhoff, B., Puistola, U., Azim, S. A., Muller-Holzner, E., Ben-Arie, A., van Kempen, L. C., Petru, E., Jahn, S., Geels, Y. P., Massuger, L. F., Amant, F., Polterauer, S., Lappi-Blanco, E., Bulten, J., Meuter, A., Tanouye, S., Oppelt, P., Stroh-Weigert, M., Reinthaller, A., Mariani, A., Hackl, W., Netzer, M., Schirmer, U., Vergote, I., Altevogt, P., Marth, C., and Fogel, M. (2013). L1CAM in Early-Stage Type I Endometrial Cancer: Results of a Large Multicenter Evaluation. *J. Natl. Cancer Inst.*, **7**, 105(15), 1142–1150. Epub 2013 Jun 18.

Histocompatibility. A patient who is waiting for an organ transplant needs a histocompatible donor who matches the patient's human leukocyte antigen (HLA) type. For a given patient, the number of matching donors per 1000 National Blood Bank records is modeled as Poisson with unknown rate λ. If a randomly selected group of 1000 records showed exactly one match, estimate λ in Bayesian fashion.

For λ assume,

(a) gamma $\mathcal{G}a(2,1)$ prior;

(b) flat prior $\lambda = 1$, for $\lambda > 0$;

(c) invariance prior $\pi(\lambda) = \frac{1}{\lambda}$, for $\lambda > 0$;

(d) Jeffreys prior $\pi(\lambda) = \frac{1}{\sqrt{\lambda}}$, for $\lambda > 0$.

Note that the priors in (b-d) are not proper densities (the integrals are not finite), however, the resulting posteriors are proper.

Neurons Fire in Potter's Lab. Data set consisting of 989 firing times in a cell culture of neurons, recorded time instances when a neuron sent a signal to another linked neuron (a spike). The cells from the cortex of an embryonic rat brain were cultured for 18 days on multielectrode arrays. The measurements were taken while the culture was stimulated at a rate of 1 Hz. From this data set the counts of firings in consecutive time intervals of length 20 milliseconds was derived:

20	19	26	20	24	21	24	29	21	17
23	21	19	23	17	30	20	20	18	16
14	17	15	25	21	16	14	18	22	25
17	25	24	18	13	12	19	17	19	19
19	23	17	17	21	15	19	15	23	22

It is believed that the counts are distributed as Poisson with unknown parameter λ. An expert believes that the number of counts in the interval of 20 milliseconds should be about 15.

(a) What is the likelihood function for these 50 observations?

(b) Using the information the expert provided elicit an appropriate Gamma prior. Is such a prior unique?

(c) For the prior suggested in (b) find Bayes' estimator of λ. How does this estimator compare to the MLE?

Elicit Inverse Gamma Prior. Specify the inverse gamma prior

$$\pi(\theta) = \frac{\beta^\alpha \exp\{-\theta/\beta\}}{\Gamma(\alpha)\theta^{\alpha+1}}, \;\; \theta \geq 0; \alpha, \beta > 0,$$

if $\mathbb{E}\theta = 2$ and $\mathbf{Var}(\theta) = 12$, are elicited from the experts.

Derive Jeffreys' Priors for Poisson λ, Bernoulli p, and Geometric p. Recall that Jeffreys' prior for parameter θ in the likelihood $f(x|\theta)$ is defined as

$$\pi(\theta) \propto |\det(\mathcal{I}(\theta))|^{1/2}$$

where, for univariate parameters,

$$\mathcal{I}(\theta) = \mathbb{E}\left[\frac{d \log f(x|\theta)}{d\theta}\right]^2 = -\mathbb{E}\left[\frac{d^2 \log f(x|\theta)}{d\theta^2}\right].$$

and expectation is taken wrt random variable $X \sim f(x|\theta)$.

(a) Show that Jeffreys' prior for Poisson distribution $f(x|\lambda) = \frac{\lambda^x}{x!}e^{-\lambda}, \lambda \geq 0$, is $\pi(\lambda) = \sqrt{1/\lambda}$.

(b) Show that Jeffreys' prior for Bernoulli distribution $f(x|p) = p^x(1-p)^{1-x},\ 0 \leq p \leq 1$, is $\pi(p) \propto \frac{1}{\sqrt{p(1-p)}}$, which is beta $\mathcal{B}e(1/2, 1/2)$ distribution (or Arcsin distribution).

(c) Show that Jeffreys' prior for Geometric distribution $f(x|p) = (1-p)^{x-1}p,\ x = 1, 2, \ldots;\ 0 \leq p \leq 1$, is $\pi(p) \propto \frac{1}{p\sqrt{1-p}}$.

Two Scenarios for the Probability of Succes. An experiment may lead to success with probability p, which is to be estimated. Two series of experiments were conducted:

(i) In the first scenario the experiment is repeated independently 10 times and the number of successes realized was 1;

(ii) In the second scenario the experiment was repeated until the success, and number of repetition was 10.

(a) The two likelihoods are Binomial and Geometric, and the moment matching estimate for probability of succes in both cases is $\hat{p} = 0.1$, however the classical inference for the two cases is different (CI, testing, etc.) Is there any difference in Bayesian inferences? Why yes or no.

(b) For any of the two scenarios find Bayes estimator of p if the prior is $\pi(p) = \frac{1}{p\sqrt{1-p}}$.

Jeffreys' Prior for Normal Precision.

The Jeffreys' prior on normal scale σ is $\pi(\sigma) = \frac{1}{\sigma}$. Consider the precision parameter $\tau = \frac{1}{\sigma^2}$.

Using the invariance property show that Jeffreys' prior for τ is $\pi(\tau) = \frac{1}{\tau}$.

Derive Jeffreys' Prior for Maxwell's θ .

(a) Show that Jeffreys' prior for Maxwell's rate parameter θ is proportional to $\frac{1}{\theta}$. Maxwell density is given by

$$f(x|\theta) = \sqrt{\frac{2}{\pi}}\ \theta^{3/2}\, x^2 \exp\left\{-\frac{1}{2}\,\theta\, x^2\right\}, \quad x \geq 0, \theta > 0.$$

(b) Show that the flat prior on $\log\theta$ is equivalent to $\frac{1}{\theta}$ prior on θ.

"Quasi" Jeffreys' Priors . Jeffreys himself often recommended priors different from Jeffreys' priors. For example, for Poisson rate λ he recommended $\pi(\lambda) \propto 1/\lambda$ instead of $\pi(\lambda) \propto \sqrt{1/\lambda}$.

For (μ, σ^2) Jeffreys recommended $\pi(\mu, \sigma^2) \propto 1 \times 1/\sigma^2$. This prior is obtained as the product of separate one-dimensional Jeffreys' priors for μ and σ^2. Show that simultaneous Jeffreys' prior for two dimensional parameter (μ, σ^2) is $\pi(\mu, \sigma^2) \propto 1 \times 1/\sigma^3$.

Haldane Prior for Binomial p . Haldane (1931)[2] suggested fully noninformative prior for binomial p as

$$\pi(p) \propto \frac{1}{p(1-p)} \quad \text{[beta } \mathcal{B}e(0,0) \text{ distribution]}$$

(a) Show that Haldane prior is equivalent to a flat prior on logit(p).

(b) Suppose $X \sim \mathcal{B}in(n, p)$ is observed. What is the posterior? What is the Bayes estimator of p.

(c) What is the predictive distribution for single future Bernoulli Y? What is the prediction for Y?

Eliciting a Normal Prior. We are eliciting a normal prior $\mathcal{N}(\mu, \sigma^2)$ from an expert who can specify percentiles. If 20th and 70th percentiles are specified as 2.7 and 4.8, respectively, how μ and σ should be elicited?

Jigsaw. An experiment with a sample of 18 nursery-school children involved the elapsed time required to put together a small jigsaw puzzle. The times were:

3.1	3.2	3.4	3.6	3.7	4.2	4.3	4.5	4.7
5.2	5.6	6.0	6.1	6.6	7.3	8.2	10.8	13.6

Assume that data are coming from normal $\mathcal{N}(\mu, \sigma^2)$ with $\sigma^2 = 8$. For parameter μ set a normal prior with mean 5 and variance 6.

(a) Find Bayes estimator and 95% credible set for population mean μ.

(b) Find posterior probability of hypothesis $H_0 : \mu \leq 5$.

(c) What is your prediction for a single future observation?

Jeremy and Poisson. Jeremy believes that normal model on his IQ test scores is not appropriate. After all, the scores are reported as integers. So he proposes a Poisson model; the scores to be modeled as Poisson,

$$y \sim \mathcal{P}oi(\theta).$$

An expert versed in GT student's intellectual abilities is asked to elicit a prior on θ. The expert elicits a gamma prior

$$\theta \sim \mathcal{G}a(30, 0.25).$$

[2]Haldane, J. B. S. (1931). A note on inverse probability. *Proceedings of the Cambridge Philosophical Society*, **28**, 55 – 61.

Jeremy gets the test and scores $y = 98$.

(a) What is the Bayes estimator of θ? Find this estimator exactly.

(b) Using WinBUGS/OpenBUGS confirm that simulations agree with the theoretical result in (a).

NPEB for p in Geometric Distribution. A geometric random variable X counts number of failures before the first success, when probability of a success is p (and a failure $1-p$). The PDF of X is

$$P(X = x) = (1-p)^x \times p,\ x = 0, 1, 2, \ldots;\quad 0 \le p \le 1.$$

We simulated a sample of size 2400 from geometric distribution with probability of success 0.32. The following (summarized) sample was obtained:

x	Freq
0	758
1	527
2	379
3	229
4	162
5	121
6	79
7	56
8	30
9	20
10	15
11	6
12	6
13	4
14	1
15	0
16	0
17	4
18	1
19	1
20	1
21+	0
Total	2400

(a) Develop Nonparametric Empirical Bayes Estimator if the prior on p is $g(p),\ 0 \le p \le 1$.

(b) Compute the empirical Bayes estimator developed in (a) on the simulated sample for different values of x.

Lifetimes and Predictive Distribution. Suppose that $T_1, \ldots, T_n$ are exponential $\mathcal{E}xp(\theta)$

lifetimes, where θ is the rate parameter. Let the prior on θ be exponential $\mathcal{E}xp(\tau)$, where τ is rate parameter, as well.

Denote with T total observed lifetime $\sum_{i=1}^{n} T_i$. Then, T is gamma $\mathcal{G}a(n, \theta)$ distributed. Show:

(a) Marginal (prior predictive) for T is $m_T(t) = \frac{n\tau t^{n-1}}{(\tau+t)^{n+1}},\ t > 0$.

(b) Posterior for θ given $T = t$ is gamma $\mathcal{G}a(n + 1, \tau + t)$.

$$\pi(\theta|y) = \frac{\theta^n (\tau + t)^{n+1}}{\Gamma(n + 1)} \exp\{-(\tau + t)\theta\}.$$

(c) Posterior predictive distribution for a new T^*, given $T = t$ is

$$f(t^*|t) = \int_0^\infty \theta \exp\{-\theta t^*\} \pi(\theta|t) d\theta = \frac{(n + 1)(\tau + t)^{n+1}}{(\tau + t + t^*)^{n+2}}.$$

(d) Expected value (wrt posterior predictive distribution) of T^* (that is, prediction for a new T^*) is

$$E(T^*|T = t) = \frac{\tau + t}{n}.$$

Normal Likelihood with Improper Priors.

Let $X_1, \ldots, X_n$ be iid normals $\mathcal{N}(\theta, \sigma^2)$, where

(a) θ is the parameter of interest, and σ^2 is known. Assume flat prior on θ,

$$\pi(\theta) = 1, \ \ -\infty < \theta < \infty.$$

Show that the posterior is

$$[\theta|X_1, \ldots, X_n] \sim \mathcal{N}\left(\overline{X}, \frac{\sigma^2}{n}\right),$$

where $\overline{X}$ is the mean of the observations.

(b) σ^2 is the parameter of interest, and θ is known. Let the prior on σ^2 be

$$\pi(\sigma^2) = \frac{1}{\sigma^2},\ \sigma^2 > 0.$$

Show that the posterior is inverse gamma,

$$[\sigma^2|X_1, \ldots, X_n] = \mathcal{IG}\left(\frac{n}{2}, \frac{\sum_{i=1}^{n}(X_i - \theta)^2}{2}\right)$$

where $\mathcal{IG}(a, b)$ stands for distribution with a density $f(y) = \frac{b^a}{\Gamma(a) y^{a+1}} e^{-b/y},\ a, b > 0,\ y \geq 0$.

Results/Solutions

Counts of Alpha Particles.

(a) $\alpha = 70$

Sum of frequencies is $n = \sum_{i=0}^{12} f_i = 2608$ and $Y = \sum_{i=0}^{12} X_i f_i = 10094$, which gives MLE for λ, $\hat{\lambda}_{MLE} = 10094/2608 = 3.8704$.

$Y \sim Pois(n\lambda)$. $\qquad \lambda \sim Gamma(\alpha, \beta)$

Likelihood $\times$ Prior $\propto \frac{(n\lambda)^y}{y!} e^{-n\lambda} \times \lambda^{\alpha-1} e^{-\beta\lambda}$

Posterior $\propto \lambda^{y+\alpha-1} e^{-(n+\beta)\lambda}$, which is $Gamma(y+\alpha, n+\beta)$.

Bayes rule $\hat{\lambda}_b = \frac{y+\alpha}{n+\beta} = \frac{10094+70}{2608+10} = 3.8824$.

```
model{
  nlambda <- n*lambda
  sumx ~ dpois(nlambda);
  lambda ~ dgamma(alpha, beta)
  }

list(n=2608, sumx= 10094, alpha=70, beta=10)

list(lambda=1)


                mean   sd  MC_error  val2.5pc median val97.5pc start sample
lambda 3.882 0.03849 3.716E-5 3.807     3.882     3.958          1001 1100000
```

FIGO.

```
total = 1021;
totalpositive = 181;  %rounded  0.177 * 1021 = 180.7170
totalnegative = 840; % as 1021 - 181
tp = 93; % true positive as rounded 181 * 0.514 = 93.0340
fp = 88; % false positives, as 181-93
fn = 24; % false negatives as rounded 840 * 0.029=24.3600
tn = 816; %true negatives, as 840-24
%
% (a) The table is
%            Rec         Not Rec    Total
% ------------------------------------------
% +          93            88             181
% -          24           816             840
% ------------------------------------------
% Tot       117           904            1021
```

```
 model{
for (i in 1:5){
  TP.rep[i] <- TP
  TN.rep[i] <- TN
  nd.rep[i] <- nd
  nc.rep[i] <- nc
  TP.rep[i] ~ dbin(se[i], nd.rep[i])
  TN.rep[i] ~ dbin(sp[i], nc.rep[i])
  }
se[1] ~ dunif(0,1)
sePhi[1] ~ dlogis(0,1)
sp[1] ~ dunif(0,1)
spPhi[1] ~ dlogis(0,1)
sePhi[2] ~ dunif(-5,5)
logit(se[2]) <- sePhi[2]
spPhi[2] ~ dunif(-5,5)
logit(sp[2]) <- spPhi[2]
se[3] ~ dbeta(0.5, 0.5)
sePhi[3] <- logit(se[3])
sp[3] ~ dbeta(0.5, 0.5)
spPhi[3] <- logit(sp[3])
sePhi[4] ~ dnorm(0, 0.5)
logit(se[4]) <- sePhi[4]
spPhi[4] ~ dnorm(0, 0.5)
logit(sp[4]) <- spPhi[4]
sePhi[5] ~ dnorm(0, 0.368)
logit(se[5]) <- sePhi[5]
spPhi[5] ~ dnorm(0, 0.368)
logit(sp[5]) <- spPhi[5]
}
list(nd=117, nc=904, TP=93, TN=816)
mean sd MC_error val2.5pc median val97.5pc start sample
se[1] 0.7899 0.03713 1.196E-4 0.7132 0.7915 0.8578 1001 100000
se[2] 0.7948 0.03703 1.041E-4 0.7175 0.7965 0.8625 1001 100000
se[3] 0.7925 0.03725 1.176E-4 0.7148 0.7942 0.8604 1001 100000
se[4] 0.7892 0.03691 1.125E-4 0.7125 0.7909 0.8568 1001 100000
se[5] 0.7907 0.03709 1.141E-4 0.7133 0.7922 0.8586 1001 100000
sp[1] 0.9017 0.009924 3.045E-5 0.8815 0.902 0.9204 1001 100000
sp[2] 0.9027 0.009796 2.984E-5 0.8826 0.903 0.9209 1001 100000
sp[3] 0.9022 0.009824 3.227E-5 0.8822 0.9025 0.9206 1001 100000
sp[4] 0.9014 0.009915 3.151E-5 0.8811 0.9017 0.9199 1001 100000
sp[5] 0.9017 0.009887 3.06E-5 0.8816 0.902 0.9202 1001 100000
```

Histocompatibility. *Hint:* In all cases (a-d), the posterior is gamma. Write the product $\frac{\lambda^1}{1!}\exp\{-\lambda\} \ \times \pi(\lambda)$ and match gamma parameters. The first part of the product is the likelihood when exactly one matching donor was observed.

Solution: Gamma $\mathcal{G}a(r, \mu)$ distribution for λ has a density

$$\pi(\lambda) = \frac{\mu^r \ \lambda^{r-1} \exp\{-\mu\lambda\}}{\Gamma(r)}, \ \lambda > 0.$$

Here $r = 2$ and $\mu = 1$, so $\pi(\lambda) = \lambda e^{-\lambda}$, since $\Gamma(2) = 1$.

The likelihood is Poisson, $f(x|\lambda) = \frac{\lambda^x}{x!}\exp\{-\lambda\}$, and since $X = 1$ is observed, the likelihood is $\lambda e^{-\lambda}$.

The posterior is proportional to the product of the likelihood and prior,

$$\lambda e^{-\lambda} \times \lambda e^{-\lambda} = \lambda^2 e^{-2\lambda}.$$

From this expression we conclude that the posterior is Gamma $\mathcal{G}a(3, 2)$. For any $Y \sim \mathcal{G}a(r, \mu)$, the mean EY is r/μ. Thus, the posterior mean is 3/2=1.5, and this is a Bayes estimator of λ. The posterior variance is $3/2^2$ and posterior standard deviation is $\sqrt{3}/2 = 0.8660$.

The supplied WinBUGS program gives the following MCMC approximation to the solution:

```
        mean sd     MC_error val2.5pc median val97.5pc start sample
lambda 1.495  0.863 0.002706 0.3107     1.332     3.609  10001 100000
```

The median is 1.332 and the 95% credible set for λ is $[-0.3107, 3.609]$.

Neurons Fire in Potter's Lab. (a) The likelihood is proportional to $\lambda^{\sum_{i=1}^{50} X_i} \ \exp\{-50\lambda\}$, where $\sum X_i = 989$ is the sum of all counts (total number of firings). The $\sum_i X_i$ is sufficient statistics here and has Poisson $\mathcal{P}oi(n\lambda)$ distribution.

(b) A gamma prior with mean 15 is not unique, for any x, $\mathcal{G}a(15x, x)$ is such a prior. However, the variances depend on x, For example for priors $\mathcal{G}a(150, 10)$, $\mathcal{G}a(15, 1)$, $\mathcal{G}a(1.5, 0.1)$, $\mathcal{G}a(0.15, 0.01)$, etc. have variances 1.5, 15, 150, 1500, etc. The variances indicate the degree of certainty of expert that the prior mean is 15. Large variances correspond to non-informative choices.

Since the sample variance of 50 observations is about 15, it is reasonable to take prior with larger variance, say $\mathcal{G}a(3, 0.2)$.

(c) Show that $\lambda|\sum_i X_i$ is gamma $\mathcal{G}a(\sum_i X_i + 3, n + 0.2)$.

Bayes estimator for λ can be represented as $w \times \bar{X} + (1-w) \times 15 =$ where $w = n/(n+0.2)$, emphasizing the fact that posterior mean is a compromise between MLE, $\bar{X}$, and prior mean, 15.

Elicit Inverse Gamma Prior. Show that the mean μ and variance σ^2 of an inverse gamma prior $\mathcal{IG}(\alpha, \beta)$ are connected with α and β as

$$\alpha = \frac{\mu^2}{\sigma^2} + 2, \quad \beta = \mu\left(\frac{\mu^2}{\sigma^2} + 1\right).$$

[**Result.** $\alpha = 7/3, \beta = 8/3$.]

Derive Jeffreys' Priors for Poisson λ, Bernoulli p, and Geometric p.

(a)

$$\begin{aligned}
\mathcal{I}(\lambda) &= \mathbb{E}\left[\frac{d}{d\lambda}\log\left(\frac{\lambda^x}{x!}e^{-\lambda}\right)\right]^2 \\
&= \mathbb{E}\left[\frac{d}{d\lambda}(\text{const} + x\log\lambda - \lambda)\right]^2 \\
&= \mathbb{E}(x/\lambda - 1)^2 = \mathbb{E}x^2/\lambda^2 - 2\mathbb{E}x/\lambda + 1
\end{aligned}$$

Since $\mathbb{E}x^2 = \mathbf{Var}x + (\mathbb{E}x)^2 = \lambda + \lambda^2$, $\mathcal{I}(\lambda) = 1/\lambda + 1 - 2 + 1 = 1/\lambda$.
Thus, $\pi(\lambda) \propto \sqrt{\frac{1}{\lambda}}$.

(b)

$$\begin{aligned}
f(x|p) &= p^x(1-p)^{1-x}, \; x = 1, 2, \dots \\
L &= \log(f(x|p)) = x\log(p) + (x-1)\log(1-p) \\
\frac{\partial L}{\partial p} &= \frac{x}{p} - \frac{1-x}{1-p} \\
\frac{\partial^2 L}{\partial p^2} &= -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}
\end{aligned}$$

Since $\mathbb{E}x = p$, the Fisher Information is

$$\mathcal{I}(p) = \frac{1}{p} + \frac{1-p}{(1-p)^2} = \frac{1}{p(1-p)}.$$

Thus, Jeffreys' prior is

$$\pi(p) = \sqrt{|\mathcal{I}(p)|} = \frac{1}{\sqrt{p(1-p)}}.$$

(c)

$$\begin{aligned}
f(x|p) &= (1-p)^{x-1}p, \; x = 0, 1 \\
L &= \log(f(x|p)) = (x-1)\log(1-p) + \log(p). \\
\frac{\partial L}{\partial p} &= \frac{1}{p} - \frac{x-1}{1-p} \\
\frac{\partial^2 L}{\partial p^2} &= -\frac{1}{p^2} - \frac{x-1}{(1-p)^2}
\end{aligned}$$

Since $\mathbb{E}x = 1/p$, the Fisher Information is

$$\mathcal{I}(p) = \frac{1}{p^2} + \frac{p^{-1}-1}{(1-p)^2} = \frac{1}{p^2} + \frac{1}{p(1-p)} = \frac{1}{p^2(1-p)}.$$

Thus, Jeffreys' prior is

$$\pi(p) = \sqrt{|\mathcal{I}(p)|} = \frac{1}{p\sqrt{1-p}}.$$

Two Scenarios for the Probability of Succes. (a) For Binomial, $\mathbb{E}X = np = 10p$, and $X = 1$, leading to $\hat{p} = 0.1$. For Geometric, $\mathbb{E}N = 1/p$ and $N = 10$, leading again to $\hat{p} = 0.1$. However, see Example 9.16 (page 413) in the Engineering Biostatistics textbook, so called Savage Disparity.

Since in both cases the likelihood is proportional to $p(1-p)^9$, Bayesian inference coincide and for a Bayesian the scenario is irrelevant, all that matters is one success and 10 trials.

(b) The posterior is proportional to $(1-p)^{17/2}$, which is beta $\mathcal{B}e(1, 19/2)$. Thus, Bayes' estimator is $\hat{p}_B = 2/19$.

Jeffreys' Prior for Normal Precision. We know that Jeffreys priot for σ is $\pi(\sigma) = \frac{1}{\sigma}$. The invariance property states that if $\tau = \tau(\sigma)$, then

$$\mathcal{I}^{1/2}(\tau) = \mathcal{I}^{1/2}(\sigma) \times \left|\frac{d\sigma}{d\tau}\right|.$$

Here $\sigma = \frac{1}{\sqrt{\tau}}$ and $\frac{d\sigma}{d\tau} = -\frac{1}{2}\tau^{-3/2}$. Thus $\pi(\tau) = \pi(\sigma) \times \left|\frac{d\sigma}{d\tau}\right| = \frac{1}{\sqrt{1/\tau}} \times \frac{1}{2}\tau^{-3/2} = \frac{1}{2\tau}$.

Since the derived prior is improper, we can drop constant 2 in the denominator and take

$$\pi(\tau) = \frac{1}{\tau},$$

as Jeffreys' prior for precision parameter τ.

Derive Jeffreys' Prior for Maxwell's θ . (a) Find the second derivative of the log likelihood. It is free of x, so the expectation is trivial. $\pi(\theta) = \frac{1}{\theta}$.

(b) Let $\phi = \log\theta$ have flat prior. Then

$$\pi(\theta) = \pi(\phi)\left|\frac{d\phi}{d\theta}\right| = 1 \times \frac{1}{\theta}.$$

"Quasi" Jeffreys' Priors . Denote $\phi = \sigma^2$. Then the normal likelihood is

$$L(\mu, \phi) = \frac{1}{\sqrt{2\pi\phi}} \exp\left\{-\frac{(x-\mu)^2}{2\phi}\right\},$$

and the log likelihood is

$$\ell(\mu, \phi) = \text{const} - \frac{1}{2}\log\phi - \frac{(x-\mu)^2}{2\phi}.$$

Then,

$$\begin{aligned} \frac{\partial \ell}{\partial \mu} &= -2\frac{x-\mu}{2\phi}\cdot(-1) = \frac{x-\mu}{\phi}, \\ \frac{\partial \ell}{\partial \phi} &= -\frac{1}{2\phi} + \frac{(x-\mu)^2}{2\phi^2}. \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 \ell}{\partial \mu^2} &= -\frac{1}{\phi}, \\ \frac{\partial^2 \ell}{\partial \mu\, \partial \phi} &= -\frac{x-\mu}{\phi^2} \\ \frac{\partial^2 \ell}{\partial \phi\, \partial \mu} &= \frac{\partial^2 \ell}{\partial \mu\, \partial \phi}, \text{ and} \\ \frac{\partial^2 \ell}{\partial \phi^2} &= \frac{1}{2\phi^2} - \frac{(x-\mu)^2}{\phi^3}. \end{aligned}$$

The Fisher Information matrix is

$$\mathcal{I} = -\mathbb{E}\begin{bmatrix} -\frac{1}{\phi} & -\frac{x-\mu}{\phi^2} \\ -\frac{x-\mu}{\phi^2} & \frac{1}{2\phi^2} - \frac{(x-\mu)^2}{\phi^3} \end{bmatrix} = \begin{bmatrix} \frac{1}{\phi} & 0 \\ 0 & \frac{1}{2\phi^2} \end{bmatrix}$$

and

$$\det(\mathcal{I}) = \frac{1}{2\phi^3}.$$

Jeffreys' prior is proportional to $|\det(\mathcal{I})|^{1/2}$

$$\pi(\mu, \phi) \propto \frac{1}{\phi^{3/2}} \propto 1 \times \frac{1}{\sigma^3}.$$

Haldane Prior for Binomial p .

(a) Let ψ be the logit of p, i.e.,

$$\psi = \log\frac{p}{1-p},$$

and assume that ψ is given a flat prior, $\psi \sim 1$. Then the prior on p is

$$\pi(p) = \pi(\psi) \times \left|\frac{d\psi(p)}{dp}\right| = 1 \times \left|\frac{1}{p/(1-p)} \cdot \frac{p(-1) - 1 \cdot (1-p)}{(1-p)^2}\right| = \frac{1}{p(1-p)}.$$

(b) The posterior is beta $\mathcal{B}e(x, n-x)$ and the Bayes estimator of p is $\frac{x}{n}$, which coincides with frequentist's $\hat{p}$.

(c) The predictive distribution for single future Bernoulli y is

$$f(y|x) = \frac{B(x+y, n+1-x-y)}{B(x, n-x)}.$$

Here, using $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ and $\Gamma(a+1) = a\Gamma(a)$, we can show

$$\begin{aligned} f(0|x) + f(1|x) &= \frac{\frac{\Gamma(x)\Gamma(n+1-x)}{\Gamma(n+1)} + \frac{\Gamma(x+1)\Gamma(n-x)}{\Gamma(n+1)}}{\frac{\Gamma(x)\Gamma(n-x)}{\Gamma(n)}} \\ &= \frac{\frac{\Gamma(x)(n-x)\Gamma(n-x)}{n\Gamma(n)} + \frac{x\Gamma(x)\Gamma(n-x)}{n\Gamma(n)}}{\frac{\Gamma(x)\Gamma(n-x)}{\Gamma(n)}} \\ &= \frac{n-x}{n} + \frac{x}{n} = 1. \end{aligned}$$

Thus, the distribution of future observation y given x success in n trials is

$y\|x$	0	1
prob	$\frac{n-x}{n}$	$\frac{x}{n}$

Note that prediction for future y is the mean of the posterior predictive distribution which is $\frac{x}{n}$. The same result is obtained when $Ey = p$ is integrated wrt posterior $\mathcal{B}e(x, n-x)$. Check this!

Eliciting a Normal Prior. If x_p is pth quantile (100%pth percentile) then $x_p = \mu + z_p\sigma$. A system of two equations with two unknowns is formed with z_p's in MATLAB/Octave as `norminv(0.20) = -0.8416` and `norminv(0.70)=0.5244`, or in R as

```
 > qnorm(0.20)
[1] -0.8416212
> qnorm(0.70)
[1] 0.5244005
```

The solution is $\mu = 3.99382 \approx 4, \quad \sigma = 1.53734$.

Jigsaw.

(a-b) Consult the Jeremy example. Note that the likelihood in normal with variance σ^2/n, since we have n observations.

(c) posterior mean.

Jeremy and Poisson.

Problem is conjugate and the posterior is gamma $\mathcal{G}a(128, 5/4)$. The posterior mean is $128 \cdot 4/5 = 102.4$ and variance $128 \cdot 16/25 = 81.92$. The posterior standard deviation is 9.0510.

The simple WinBUGS code is

```
model{
   X ~ dpois(lambda)
  lambda ~ dgamma(30, 0.25)
  }

DATA
list(X=98)

          mean sd MC_error val2.5pc median val97.5pc start sample
lambda 102.4 9.085 0.02726 85.42 102.1 121.0 1001 100000
```

NPEB for p in Geometric Distribution.

Solution for (a):

The likelihood and prior are

$$f(x|p) = (1-p)^x p,\ x = 0, 1, 2, \dots; \quad p \sim g(p), \quad 0 \le p \le 1;$$

leading to the marginal for X

$$m(x) = \int_0^1 f(x|p)dG(p) = \int_0^1 (1-p)^x p\ g(p)\ dp.$$

The posterior mean is

$$\begin{aligned} E(p|x) &= \frac{1}{m(x)} \int_0^1 p(1-p)^x p\ g(p)\ dp \\ &= \frac{1}{m(x)} \int_0^1 [1-(1-p)](1-p)^x p\ g(p)\ dp \\ &= \frac{1}{m(x)} \left(m(x) - \int_0^1 (1-p)^{x+1} p\ g(p)\ dp \right) \\ &= 1 - \frac{m(x+1)}{m(x)} \end{aligned}$$

An automatic estimator for $m(x)$ is

$$\hat{m}(x) = \frac{\text{\# of observations} = x}{\text{Total \# of observations}}.$$

This leads to

$$\begin{aligned} p^* &= 1 - \frac{\text{\# of observations} = x+1}{\text{\# of observations} = x}, \\ \hat{p} &= min\{1, max\{0, p^*\}\}. \end{aligned}$$

In this case $\hat{p}$ is free of prior distribution g (although the marginal depends on g).

Solution for (b): For the simulated data the estimators (at particular values of x) are given in the following table:

x	Freq	$\hat{p}$
0	758	0.3047
1	527	0.2808
2	379	0.3958
3	229	0.2926
4	162	0.2531
5	121	0.3471
6	79	0.2911
7	56	0.4643
8	30	0.3333
9	20	0.2500
10	15	0.6000
11	6	0.0000
12	6	0.3333
13	4	0.7500
14	1	1.0000
15	0	NaN
16	0	0
17	4	0.7500
18	1	0
19	1	0
20	1	1
21+	0	NaN
Total	2400	

Note that for $x \geq 10$ the NPEB estimators become unreliable due to low frequency counts.