

Math 341 / 650 Spring 2019

Midterm Examination ~~One~~ *Solutions*

Two

Professor Adam Kapelner

Tuesday, April 16, 2019

Full Name _____

Code of Academic Integrity

Since the college is an academic community, its fundamental purpose is the pursuit of knowledge. Essential to the success of this educational mission is a commitment to the principles of academic integrity. Every member of the college community is responsible for upholding the highest standards of honesty at all times. Students, as members of the community, are also responsible for adhering to the principles and spirit of the following Code of Academic Integrity.

Activities that have the effect or intention of interfering with education, pursuit of knowledge, or fair evaluation of a student's performance are prohibited. Examples of such activities include but are not limited to the following definitions:

Cheating Using or attempting to use unauthorized assistance, material, or study aids in examinations or other academic work or preventing, or attempting to prevent, another from using authorized assistance, material, or study aids. Example: using an unauthorized cheat sheet in a quiz or exam, altering a graded exam and resubmitting it for a better grade, etc.

I acknowledge and agree to uphold this Code of Academic Integrity.

signature

date

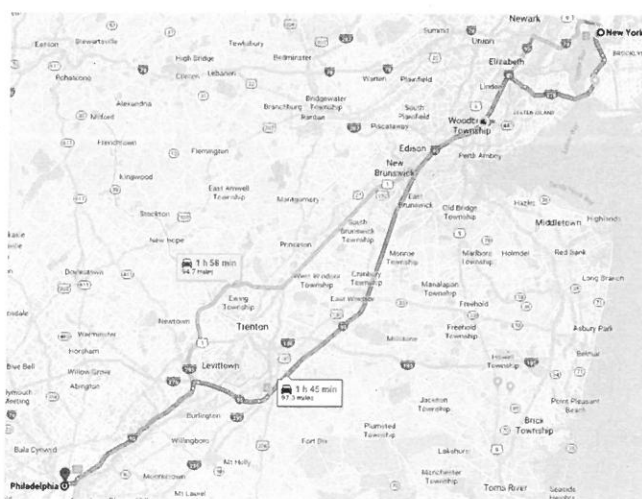
Instructions

This exam is seventy five minutes and closed-book. You are allowed **one** page (front and back) of a "cheat sheet." You may use a graphing calculator of your choice. Please read the questions carefully. If the question reads "compute," this means the solution will be a number otherwise you can leave the answer in *any* widely accepted mathematical notation which could be resolved to an exact or approximate number with the use of a computer. I advise you to skip problems marked "[Extra Credit]" until you have finished the other questions on the exam, then loop back and plug in all the holes. I also advise you to use pencil. The exam is 100 points total plus extra credit. Partial credit will be granted for incomplete answers on most of the questions. Box in your final answers. Good luck!

Distribution of r.v.	Quantile Function	PMF / PDF function	CDF function	Sampling Function
beta	$qbeta(p, \alpha, \beta)$	$d-(x, \alpha, \beta)$	$p-(x, \alpha, \beta)$	$r-(\alpha, \beta)$
betabinomial	$qbetabinom(p, n, \alpha, \beta)$	$d-(x, n, \alpha, \beta)$	$p-(x, n, \alpha, \beta)$	$r-(n, \alpha, \beta)$
binomial	$qbinom(p, n, \theta)$	$d-(x, n, \theta)$	$p-(x, n, \theta)$	$r-(n, \theta)$
exponential	$qexp(p, \theta)$	$d-(x, \theta)$	$p-(x, \theta)$	$r-(\theta)$
gamma	$qgamma(p, \alpha, \beta)$	$d-(x, \alpha, \beta)$	$p-(x, \alpha, \beta)$	$r-(\alpha, \beta)$
inversegamma	$qinvgamma(p, \alpha, \beta)$	$d-(x, \alpha, \beta)$	$p-(x, \alpha, \beta)$	$r-(\alpha, \beta)$
negative-binomial	$qnbinom(p, r, \theta)$	$d-(x, r, \theta)$	$p-(x, r, \theta)$	$r-(r, \theta)$
normal (univariate)	$qnorm(p, \theta, \sigma)$	$d-(x, \theta, \sigma)$	$p-(x, \theta, \sigma)$	$r-(\theta, \sigma)$
poisson	$qpois(p, \theta)$	$d-(x, \theta)$	$p-(x, \theta)$	$r-(\theta)$
T (standard)	$qt(p, \nu)$	$d-(x, \nu)$	$p-(x, \nu)$	$r-(\nu)$
T (nonstandard)	$qt.scaled(p, \nu, \mu, \sigma)$	$d-(x, \nu, \mu, \sigma)$	$p-(x, \nu, \mu, \sigma)$	$r-(\nu, \mu, \sigma)$
uniform	$qunif(p, a, b)$	$d-(x, a, b)$	$p-(x, a, b)$	$r-(a, b)$

Table 1: Functions from R (in alphabetical order) that can be used on assignments and exams. The hyphen in columns 3, 4 and 5 is shorthand notation for the full text of the r.v. which can be found in column 2.

Problem 1 Uber is evaluating a new pool product for intercity ridesharing. They are first trying to estimate the demand in number of riders θ going from New York City to Philadelphia per weekday.



To estimate this demand, they have piloted the program among their VIP customers and collected the number of people who requested such pooled rides for $n = 20$ weekdays. The data is below:

39 50 57 50 42 50 56 68 49 50 53 53 48 44 56 30 57 49 54 48

Summary statistics for this data are: $\bar{x} = 50.15$ and $s = 7.809$.

- (a) [6 pt / 6 pts] In modeling this dataset, assume $\mathcal{F} : X_1, \dots, X_{20} \mid \theta \stackrel{iid}{\sim} \text{Poisson}(\theta)$. Explain why this assumption is reasonable.

The Poisson model for # of riders on any given day is reasonable because there are a very large number of people in NYC but each have a very small probability of needing to travel to Philadelphia on any given day. The iid assumption also is not unreasonable because each day's ridership is independent of previous days' ridership and the average mean number of riders (θ) is likely constant.

- (b) [6 pt / 12 pts] Find the posterior distribution (including the values of the hyperparameters to two decimal places). List all assumptions explicitly.

Assume: $P(\theta) = \text{Gamma}(\alpha, \beta)$

$$\Rightarrow P(\theta | x) = \text{Gamma}(\sum x_i + \alpha, n + \beta) = \text{Gamma}(1003, 20)$$

$$\sum x_i = \bar{x}n = 50.15 \cdot 20 = 1003$$

- (c) [6 pt / 18 pts] Using your answer from (b), provide a Bayesian point estimate of θ .

$$\hat{\theta}_{\text{MTE}} = \frac{\sum x_i}{n} = \bar{x} = 50.15$$

- (d) [6 pt / 24 pts] Uber will only turn a profit if there are 55 or more riders going from NYC to Philadelphia per weekday. Find the probability they will turn a profit using your answer from (b).

$$P(\theta \geq 55 | x) = 1 - \text{pgamma}(55, 1003, 20)$$

- (e) [6 pt / 30 pts] Find the probability there are 55 or more riders going from NYC to Philadelphia tomorrow (Wednesday) using the data and whatever assumptions you made in (b).

$$P(x^* \geq 55 | x) = 1 - \text{pbinom}(54, 1003, 0.952)$$

$$\begin{aligned} P(x^* | x) &= \text{Negbin}\left(\sum x_i, \frac{n}{n+1}\right) \\ &= \text{Negbin}\left(1003, \frac{20}{21} = 0.952\right) \end{aligned}$$

Problem 2 Consider the geometric model $\mathcal{F} : X_1, \dots, X_n \mid \theta \stackrel{iid}{\sim} \text{Geom}(\theta) := (1 - \theta)^x \theta$, its PMF, $\theta \in \Theta = (0, 1)$, $\text{Supp}[X] = \{0, 1, 2, \dots\}$, $\mathbb{E}[X] = \frac{1-\theta}{\theta}$ and $\text{Var}[X] = \frac{1-\theta}{\theta^2}$.

(a) [6 pt / 36 pts] Find $\mathbb{P}(X \mid \theta)$ where X refers to all data (i.e. X_1, \dots, X_n). Simplify.

$$P(X|\theta) = \prod_{i=1}^n (1-\theta)^{x_i} \theta = \theta^n (1-\theta)^{\sum x_i}$$

(b) [6 pt / 42 pts] Show that the beta distribution is conjugate for the geometric likelihood.

$$\text{Assume: } P(\theta) = \text{Beta}(\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

$$P(\theta|x) \propto P(X|\theta) P(\theta) = \left(\theta^n (1-\theta)^{\sum x_i} \right) \left(\frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} \right) \propto \theta^{n+\alpha-1} (1-\theta)^{\sum x_i + \beta - 1} \\ \propto \text{Beta}(n+\alpha, \sum x_i + \beta)$$

Since the posterior is beta when the prior is beta \Rightarrow beta is the conjugate prior for the geometric likelihood model

(c) [9 pt / 51 pts] Show that Jeffrey's Prior is a beta distribution.

$$\ell(\theta; x) = \ln(P(X|\theta)) = n \ln(\theta) + \sum x_i \ln(1-\theta)$$

$$\ell'(\theta; x) = \frac{n}{\theta} - \frac{\sum x_i}{1-\theta}$$

$$\ell''(\theta; x) = -\frac{n}{\theta^2} + \frac{\sum x_i}{(1-\theta)^2}$$

$$-\ell''(\theta; x) = \frac{n}{\theta^2} + \frac{\sum x_i}{(1-\theta)^2}$$

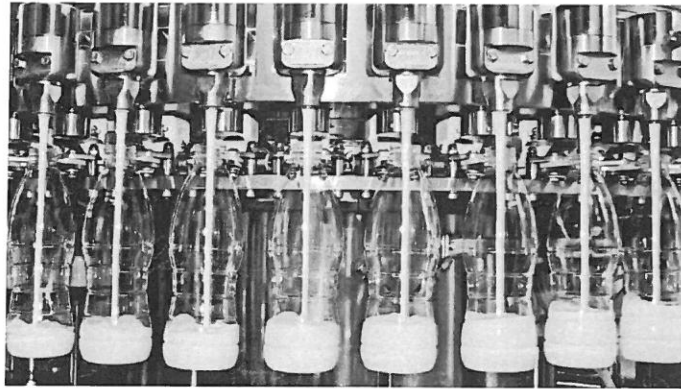
$$P(\theta) \propto \sqrt{I(\theta)} = \sqrt{n \theta^2 (1-\theta)^{-1}} \propto \theta^{-1} (1-\theta)^{-\frac{1}{2}} \\ = \theta^{(-1)} (1-\theta)^{(\frac{1}{2})-1} \propto \text{Beta}\left(0, \frac{1}{2}\right)$$

$$\theta = \mathbb{E}_x[-\ell''(\theta; x)] = \frac{n}{\theta^2} + \frac{n \mathbb{E}_x(X)}{(1-\theta)^2} = \frac{n}{\theta^2} + \frac{n \left(\frac{1-\theta}{\theta}\right)}{(1-\theta)^2} = \frac{n}{\theta^2} + \frac{n}{\theta(1-\theta)} = n \left(\frac{1-\theta}{\theta^2(1-\theta)} + \frac{\theta}{\theta^2(1-\theta)} \right) = n \theta^{-2} (1-\theta)^{-1}$$

(d) [1 pt / 52 pts] Is Jeffrey's prior you found in the previous question proper? Yes / no.

- (e) [5 pt / 57 pts] [Extra Credit] Find $\mathbb{P}(X^* | X)$ for one new sample (i.e. $n^* = 1$) and simplify.

Problem 3 This question is about building a model to understand the accuracy of this beverage-filling machine



which fills 12oz plastic bottles. We decide to do an experiment and select $n = 50$ bottles at random and measure the amount of liquid in each bottle. Here are the volumes (in oz):

10.76	11.02	11.62	10.20	12.03	12.18	12.06	11.42	11.93	10.68
11.81	11.27	11.68	11.31	11.29	12.37	12.00	10.74	12.04	10.69
11.82	11.50	11.48	11.81	11.24	11.67	11.02	11.52	11.98	11.99
11.99	11.85	11.42	11.95	10.77	11.56	11.77	11.43	12.03	11.89
11.50	11.75	11.30	12.26	12.22	12.94	10.80	12.01	11.61	12.35

and sample statistics: $\bar{x} = 11.61\text{oz}$ and $s = 0.533\text{oz}$. We will consider these measurements realizations from the model $\mathcal{F} : X_1, \dots, X_n \mid \theta, \sigma^2 \stackrel{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$.

- (a) [8 pt / 65 pts] The company is interested in testing if the bottles are being underfilled (i.e. less than 12oz). Even though it is unrealistic, assume that $\sigma^2 = s^2 = 0.533^2 = 0.284$. Use an uninformative prior and set the significant level of the test to be $\alpha = 5\%$. Calculate as exactly as you can.

$$p(\theta) \propto 1 \Rightarrow p(\theta|x) = N(\bar{x}, \frac{\sigma^2}{n}) = N(11.61, \frac{0.533^2}{50}) = N(11.61, 0.0572)$$

$$H_0: \theta \geq 12, H_A: \theta < 12, \alpha = 5\%$$

$$p_{val} = 1 - \text{pnorm}(12, 11.61, 0.0572) = P(Z \geq \frac{12 - 11.61}{0.0572}) = P(Z \geq 5.2) \approx 0 < \alpha = 5\%$$

\Rightarrow Reject $H_0 \Rightarrow$ the bottles are being underfilled!

- (b) [6 pt / 71 pts] What is the probability a bottle selected at random will be underfilled? Leave your answer in terms of the `pnorm` function with arguments rounded to two decimal places. Do not solve explicitly.

$$P(X^* \leq 12 | X, \sigma^2) = \text{pnorm}(12, 11.61, \sqrt{0.533^2 + 0.0572}) = \text{pnorm}(12, 11.61, 0.54)$$

$$p(x^* | x, \sigma^2) = N(\bar{x}, \sigma^2 + \sigma_p^2)$$

- (c) [6 pt / 77 pts] Design a prior centered at 12oz with the strength of 100 samples. Note: this is a standalone question.

$$p(\theta | \sigma^2) = N(\mu_0, \frac{\sigma^2}{n_0}) = N(12, \frac{0.533^2}{100}) = N(12, 0.00284) \\ = N(12, 0.0533^2)$$

- (d) [6 pt / 83 pts] The company is interested understanding if the bottles are being filled too erratically. Even though it is unrealistic, assume the machine is filling accurately on average (i.e. $\theta = 12\text{oz}$ exactly). Find a Bayesian estimate of σ^2 explicitly to two decimal places (an answer using a function in Table 1 is not acceptable here). I have precomputed $\hat{\sigma}_{MLE}^2 = \frac{1}{50} \sum_{i=1}^{50} (x_i - 12)^2 = 0.4305$. List your assumptions explicitly.

$$\text{Assume: } p(\sigma^2) = \text{InvGamma}(0, 0) \Rightarrow p(\sigma^2 | x, 0) = \text{InvGamma}(\frac{n}{2}, \frac{n \hat{\sigma}_{MLE}^2}{2}) \\ = \text{InvGamma}(\frac{50}{2}, \frac{50 \cdot 0.4305}{2}) \\ = \text{InvGamma}(25, 10.76)$$

$$\hat{\sigma}_{MSE}^2 = \frac{10.76}{25-1} = 0.45$$

- (e) [6 pt / 89 pts] Given the assumptions from the previous question of mean being known and an uninformative prior for σ^2 , find the probability a bottle selected at random will be overfilled (i.e. 14oz or more).

$$P(X^* \geq 14 | X, \theta) = 1 - \text{pt.scaled}(14, 50, 12, 0.43)$$

$$P(X^* | X, \theta) = T_h(\theta, \hat{\sigma}_{MLE}^2) = T_{50}(12, 0.43)$$

Problem 4 Below are theoretical questions that are independent of each other.

- (a) [6 pt / 95 pts] If $X | \theta, n \sim \text{Binomial}(n, \theta)$ where n is known, the prior on theta is Beta(α, β) where α and β are both > 0 and non-integer and X^* is the number of successes in the next n^* trials, find the kernel of the posterior predictive distribution, i.e. $k(X^* | X, n)$. Simplify as much as possible.

$$P(X^* | X, n) = \text{Beta Binom}(n^*, X + \alpha, n - X + \beta) = \binom{n^*}{X^*} \frac{B(X^* + X + \alpha, n^* - X^* + n - X + \beta)}{B(X + \alpha, n - X + \beta)}$$

$$\propto \binom{n^*}{X^*} B(X^* + X + \alpha, n^* - X^* + n - X + \beta) = \frac{n^*!}{X^*!(n^* - X^*)!} \frac{\Gamma(X + \alpha) \Gamma(n - X + \beta)}{\Gamma(n^* + \alpha + \beta)}$$

$$\propto \frac{\Gamma(X^* + X + \alpha) \Gamma(n^* - X^* + n - X + \beta)}{X^*!(n^* - X^*)!}$$

(b) [6 pt / 101 pts]

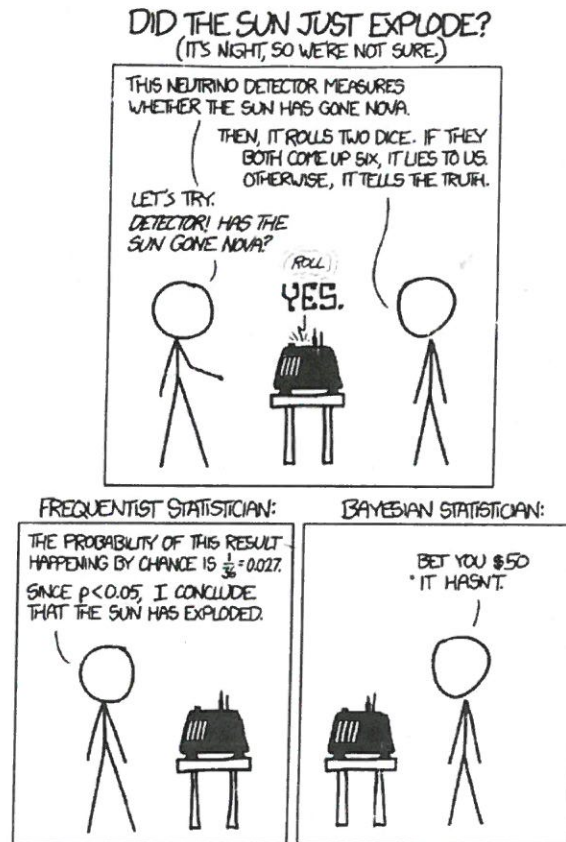
Compute the following integral as a function of a , b and fundamental constants:

$$\int_{\mathbb{R}} e^{ax-bx^2} dx = \sqrt{\frac{\pi}{b}} e^{\frac{a^2}{4b}} \int_{\mathbb{R}} \sqrt{\frac{b}{\pi}} e^{-\frac{b}{4b} \left(x - \frac{a}{2b}\right)^2} e^{ax-bx^2} dx = \sqrt{\frac{\pi}{b}} e^{\frac{a^2}{4b}}$$

$$e^{ax-bx^2} \propto N\left(\frac{a}{2b}, \frac{1}{2b}\right) = \frac{1}{\sqrt{2\pi \frac{1}{2b}}} e^{-\frac{1}{2 \cdot \frac{1}{2b}} \left(x - \frac{a}{2b}\right)^2} = \sqrt{\frac{b}{\pi}} e^{-b\left(x^2 - \frac{2ax}{2b} + \frac{a^2}{4b^2}\right)}$$

$$= \sqrt{\frac{b}{\pi}} e^{-bx^2 + ax - \frac{a^2}{4b}} = \sqrt{\frac{b}{\pi}} e^{-\frac{a^2}{4b}} e^{ax-bx^2}$$

(c) [4 pt / 105 pts]



Choose the best sentence completion. "The comic above is an example of ..."

(I) normal-normal conjugacy (II) using an informative prior (III) using the uninformative Jeffrey's prior (IV) compound distribution (V) point estimation using $\hat{\theta}_{MLE}$.