Math 341 / 650 Fall 2021
Midterm Examination One

Professor Adam Kapelner

Wednesday, March 10, 2021

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.

By taking this exam, you acknowledge and agree to uphold this Code of Academic Integrity.

Instructions

This exam is 70 minutes (variable time per question) and closed-book. You are allowed **one** page (front and back) of a "cheat sheet", blank scrap paper and a graphing calculator. Please read the questions carefully. No food is allowed, only drinks.

Problem 1 [7min] (and 7min will have elapsed) These are conceptual questions about statistical inference in the Bayesian and Frequentist perspectives.

- [10 pt / 10 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) The goal of point estimation is to find an approximation of the true value of θ.

  (b) The goal of theory testing is to find an approximation of the true value of θ.

  (c) The goal of statistical inference (in general) is to learn the true values of $x_1, \ldots, x_n$.

  (d) A parametric model $\mathcal{F}$ must be assumed to do statistical inference using the Frequentist perspective.

  (e) A parametric model $\mathcal{F}$ must be assumed to do statistical inference using the Bayesian perspective.

  (f) The data $x_1, \ldots, x_n$ cannot be collected without assuming an $\mathcal{F}$.

  (g) To compute the value of the likelihood for the data $x_1, \ldots, x_n$, you must assume an $\mathcal{F}$.

  (h) In the Frequentist perspective, $\mathbb{P}(\theta)$ is degenerate.

  (i) In the Bayesian perspective, you can only do inference if the data is realized from $\overset{iid}{\sim}$ random variables.

  (j) In the Bayesian perspective, to compute the posterior, you must assume $\mathcal{F}$ and $\mathbb{P}(\theta)$.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 2 [9min] (and 16min will have elapsed) Assume $X_1, \dots, X_n \stackrel{iid}{\sim} p(x;\theta)$, a discrete rv with support $\mathcal{X}$ and parameter space Θ and $p(\theta)$ is the PMF for a rv with support Θ.

- [14 pt / 24 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) $p(x_1, \dots, x_n; \theta) = p(x_1;\theta) \cdot p(x_2;\theta) \cdot \ldots \cdot p(x_n;\theta)$

(b) $\mathcal{L}(\theta; x_1, \dots, x_n) = \mathcal{L}(\theta; x_1) \cdot \mathcal{L}(\theta; x_2) \cdot \ldots \cdot \mathcal{L}(\theta; x_n)$

(c) $\ell(\theta; x_1, \dots, x_n) = \ell(\theta; x_1) \cdot \ell(\theta; x_2) \cdot \ldots \cdot \ell(\theta; x_n)$

(d) You can compute the maximum likelihood estimate by setting $\sum_{i=1}^n \dfrac{d\ell(\theta; x_i)}{d\theta} = 0$ and solving for θ.

(e) You can compute the maximum likelihood estimate by setting $\sum_{i=1}^n \dfrac{d\ell(\theta; x_i)}{dx_i} = 0$ and solving for θ.

(f) If n is large, the maximum likelihood estimator is approximately normally distributed.

(g) $\sum_{x\in\mathcal{X}} p(x;\theta) = 1$.

(h) $\sum_{x\in\mathcal{X}} p(\theta;x) = 1$.

(i) $\sum_{x\in\mathcal{X}} p(x) = 1$.

(j) $\sum_{x\in\mathcal{X}} p(\theta) = 1$.

(k) $\sum_{\theta\in\Theta} p(x;\theta) = 1$.

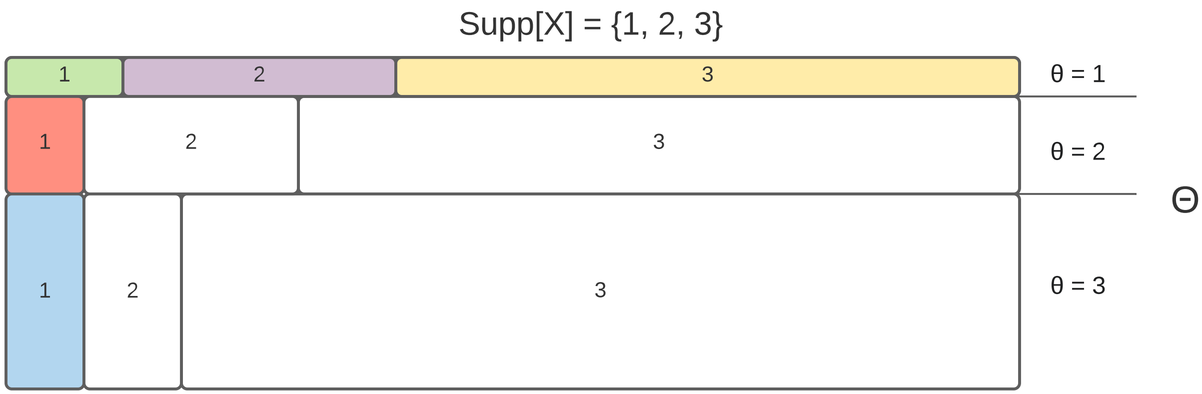
(l) $\sum_{\theta\in\Theta} p(\theta;x) = 1$.

(m) $\sum_{\theta\in\Theta} p(x) = 1$.

(n) $\sum_{\theta\in\Theta} p(\theta) = 1$.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 3 [9min] (and 25min will have elapsed) Consider a rv X with parameter space Θ. Below is an illustration of the universe of all values of x and θ drawn to-scale. The values inside the boxes are values of $x \in \mathbb{S}\text{upp}[X]$ and the values in the right margin are the values of $\theta \in \Theta$.



- [9 pt / 33 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) X is a discrete rv.

  (b) Θ is discrete.

  (c) If you add up the areas of all the boxes above, you will get 100%.

  (d) The prior was created using the principle of indifference.

  (e) $\mathbb{P}(X = 1, \theta = 1) > \mathbb{P}(X = 1, \theta = 3)$.

  (f) $\mathbb{P}(X = 1 \mid \theta = 1) > \mathbb{P}(X = 1 \mid \theta = 3)$.

  (g) $\mathbb{P}(\theta = 1 \mid X = 3)$ can be computed by taking the yellow area and dividing by the sum of the yellow area, purple area and green area.

  (h) $\mathbb{P}(\theta = 3 \mid X = 1)$ can be computed by taking the blue area and dividing by the sum of the blue area, red area and green area.

  (i) $\mathbb{P}(\theta = 1 \mid X = 1)$ can be computed by taking the green area and dividing by the sum of the green area, purple area and yellow area.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 4 [8min] (and 33min will have elapsed) Assume $X_1, X_2, X_3, X_4 \overset{iid}{\sim}$ Bernoulli(θ) and $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0$. For inference performed from the Bayesian perspective, assume Laplace's prior of indifference.

- [7 pt / 40 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) The maximum likelihood estimate of θ is $\bar{x}$.

(b) The maximum likelihood estimate of θ is $\frac{1}{2}$.

(c) The maximum likelihood estimate of θ is 0.

(d) $CI_{\theta,90\%} = \{0\}$.

(e) $CI_{\theta,100\%} = \mathbb{R}$.

(f) A frequentist hypothesis test of $H_a : \theta \neq 0.01$ at significance level 0.01 has a retainment region of $\{0\}$.

(g) A frequentist hypothesis test of $H_a : \theta \neq 0.01$ at significance level 0.01 will result in a rejection of H_0.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 5 [11min] (and 44min will have elapsed) Assume $X_1, X_2, X_3, X_4 \overset{iid}{\sim} \text{Bernoulli}(\theta)$ and $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0$. For inference performed from the Bayesian perspective, assume Laplace's prior of indifference.

- [16 pt / 56 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) The posterior after seeing only x_1 is $\mathbb{P}(\theta \mid x) = \text{Beta}(1,\, 1)$.

(b) The posterior after seeing only x_1 is $\mathbb{P}(\theta \mid x) = \text{Beta}(1,\, 2)$.

(c) The full posterior after all $n = 4$ observations is $\mathbb{P}(\theta \mid x) = \text{Beta}(1,\, 5)$.

(d) The full posterior after all $n = 4$ observations is $\mathbb{P}(\theta \mid x) = \text{Beta}(5,\, 1)$.

(e) The prior predictive distribution is $\mathbb{P}(x) = \text{Beta}(1,\, 1)$.

(f) The prior predictive distribution is $\mathbb{P}(x) = \text{Binomial}(n,\, \theta)$.

(g) The maximum a posteriori estimate of θ is 0.

(h) The maximum a posteriori estimate of θ is > 0.

(i) The minimum mean squared error estimate of θ is 0.

(j) The minimum mean squared error estimate of θ is $1/5$.

(k) The minimum mean squared error estimate of θ is $1/6$.

(l) The prior expectation is $1/2$.

(m) The prior median is $1/2$.

(n) The only prior mode is $1/2$.

(o) The minimum mean squared error estimator of θ can be written as $\dfrac{2}{3}\,\hat{\theta}^{\text{MLE}} + \dfrac{1}{3}\,\mathbb{E}[\theta]$.

(p) The $\hat{\theta}_{\text{MMSE}}$ and $\hat{\theta}_{\text{MMAE}}$ estimators for the value of θ do not have any shrinkage whatsoever.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 6 [11min] (and 55min will have elapsed) Assume $X_1, X_2, X_3, X_4 \stackrel{iid}{\sim}$ Bernoulli(θ) and $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0$. For inference performed from the Bayesian perspective, assume Laplace's prior of indifference.

- [13 pt / 69 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) $CR_{\theta,90\%} = [\text{qbeta}(5\%, 1, 5), \text{qbeta}(95\%, 1, 5)]$

(b) $CR_{\theta,90\%} = [0, \text{qbeta}(90\%, 1, 5)]$

(c) $CR_{\theta,90\%} = [\text{qbeta}(10\%, 1, 5), 1]$

(d) One needs to declare the test's significance level in order to compute a Bayesian p-value

(e) When testing $H_a : \theta < 0.3$, the Bayesian p-value can be computed via pbeta(0.3, 1, 5).

(f) When testing $H_a : \theta < 0.3$, the Bayesian p-value can be computed via pbeta(0.7, 1, 5).

(g) When testing $H_a : \theta < 0.3$, the Bayesian p-value can be computed via 1 - pbeta(0.3, 1, 5).

(h) When testing $H_a : \theta < 0.3$, the Bayesian p-value can be computed via 1 - pbeta(0.7, 1, 5).

(i) When testing $H_a : \theta < 0.3$, the Bayesian p-value can be computed via $\int_{0.3}^{1} \frac{1}{B(1,5)}(1-\theta)^4 d\theta$.

(j) When testing $H_a : \theta < 0.3$, the Bayesian p-value can be computed via $\int_{0}^{0.3} \frac{1}{B(1,5)}(1-\theta)^4 d\theta$.

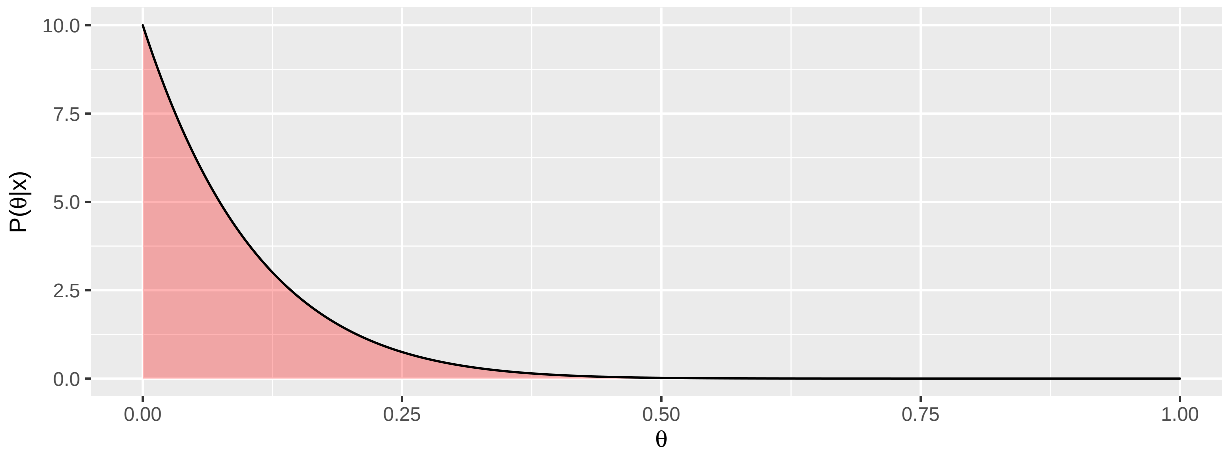
(k) When testing $H_a : \theta < 0.3$, the Bayesian p-value can be computed via $\int_{0}^{1} \frac{1}{B(1,5)}(1-\theta)^4 d\theta$.

(l) When testing $H_a : \theta \neq 0.3$, the Bayesian p-value is zero.

(m) When testing $H_a : \theta \notin [0.3 \pm \delta]$, the Bayesian p-value is zero for all $\delta > 0$.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.

Problem 7 [15min] (and 70min will have elapsed) This question is independet of those that came previously. Assuming a parametric model of the binomial with fixed n and a beta prior, the posterior is plotted below:



- [11 pt / 80 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) $HDR_{\theta,90\%} = [a, b]$ where the values a and b satisfies $0 < a < b < 1$.

  (b) $HDR_{\theta,90\%} = [0, b]$ where the value b satisfies $0 < b < 1$.

  (c) If the Haldane prior was chosen, then we can be certain that $x > 0$ and $x < n$.

  (d) If the Haldane prior was chosen, there is no shrinkage in $\hat{\theta}_{\text{MMSE}}$.

  (e) $\hat{\theta}_{\text{MAP}} = \hat{\theta}_{\text{MMAE}}$.

  (f) $\hat{\theta}_{\text{MMAE}} = \hat{\theta}_{\text{MMSE}}$.

  (g) $\hat{\theta}_{\text{MMAE}} = 1 - (1/2)^{1/10}$.

  (h) $\hat{\theta}_{\text{MMAE}} = 0.5$.

  (i) When testing $H_a : \theta < 0.3$, the Bayesian p-value is 0.0282 to the nearest three significant digits.

  (j) When testing $H_a : \theta < 0.3$, the Bayesian p-value is 0.282 to the nearest three significant digits.

Your answer will consist of a lowercase string (e.g. `aebgd`) where the order of the letters does not matter.