Math 369 / 690 Fall 2021
Final Examination

Professor Adam Kapelner

Thursday, December 16, 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 110 minutes (variable time per question) and closed-book. You are allowed **three** page (front and back) of a "cheat sheet", blank scrap paper and a graphing calculator. Please read the questions carefully. I recommend answering all questions that are easy first and then circling back to work on the harder ones. Gray text means the text is repeated verbatim from a previous problem. No food is allowed, only drinks.

Problem 1 [13min] (and 13min will have elapsed) Benford's law discovered in 1938 a discrete parameterless probability distribution on the first digits of numbers. It is also called the "first-digit law" and was discovered by examining real-world datasets. In real-world datasets, numbers beginning with a 1 are about twice as likely as numbers beginning with a 2. Numbers beginning with a 2 are about 50% more likely that numbers beginning with a 3, etc. This is especially true in domains such as accounting and taxes. This distribution is used frequently to prove that a dataset is fraudulent i.e. if someone is trying to cheat on their taxes by randomly generating numeric values, their values will not follow Benford's law since most crooks aren't too careful about learning probability theory. Below is the Benford Law distribution's PMF and CDF. It's mean is 3.441.

x	1	2	3	4	5	6	7	8	9
$p(x)$	.301	.176	.125	.097	.079	.067	.058	.051	.046
$F(x)$	.301	.477	.602	.699	.778	.845	.903	.954	1.000

Consider a tax return with $n = 45$ numbers. We examine the first digit of the numbers and sort the data. It turns out the first digit is exactly uniformly distributed across all digits i.e. $\boldsymbol{x} = < 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, \ldots, 9, 9, 9, 9, 9 >$. This smells of fraud; we will investigate. To test cheating we can use ...

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

(a) ... the one sample asymptotic z test for the mean
(b) ... the one sample asymptotic t test for the mean
(c) ... the one sample Wald test for the mean
(d) ... the score test for the mean
(e) ... the likelihood ratio test for the mean
(f) ... a generalized likelihood ratio test
(g) ... Pearson's χ^2 goodness of fit test
(h) ... Kolmogorov-Smirnov's one-sample test
(i) ... a bootstrap test where $\bar{x}$ is calculated for each bootstrap sample
(j) ... the Welch-Satterthwaite t test
(k) ... Kolmogorov-Smirnov's two-sample test
(l) ... Fisher's two-sample permutation test
(m) ... the multivariate delta method

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

Problem 2 [14min] (and 27min will have elapsed) Benford's law discovered in 1938 a discrete parameterless probability distribution on the first digits of numbers. It is also called the "first-digit law" and was discovered by examining real-world datasets. In real-world datasets, numbers beginning with a 1 are about twice as likely as numbers beginning with a 2. Numbers beginning with a 2 are about 50% more likely that numbers beginning with a 3, etc. This is especially true in domains such as accounting and taxes. This distribution is used frequently to prove that a dataset is fraudulent i.e. if someone is trying to cheat on their taxes by randomly generating numeric values, their values will not follow Benford's law since most crooks aren't too careful about learning probability theory. Below is the distribution's PMF and CDF. It's mean is 3.441.

x	1	2	3	4	5	6	7	8	9
$p(x)$	.301	.176	.125	.097	.079	.067	.058	.051	.046
$F(x)$	.301	.477	.602	.699	.778	.845	.903	.954	1.000

Consider a tax return with $n = 45$ numbers. We examine the first digit of the numbers and sort the data. It turns out the first digit is exactly uniformly distributed across all digits i.e. $\boldsymbol{x} =< 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, \ldots, 9, 9, 9, 9, 9 >$. This smells of fraud; we will investigate. To do so, we will employ Pearson's χ^2 goodness of fit test at $\alpha = 5\%$. Critical values you may need to reference are: $F_{\chi^2_7}(14.1) = F_{\chi^2_8}(15.5) = F_{\chi^2_9}(16.9) = F_{\chi^2_{10}}(18.3) = .95$.

- [12 pt / 25 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) $H_0 : \theta = 3.441$ where θ denotes the mean of the DGP that generated $x_1, \ldots, x_{45}$

(b) H_0 : the DGP that generated $x_1, \ldots, x_{45}$ is Benford's law

(c) H_0 : the PMF of the DGP that generated $x_1, \ldots, x_{45}$ is the $p(x)$ given as $p(x)$ in the table above

(d) Pearson's χ^2 goodness of fit test statistic is a realization from an approximate χ^2_7 distribution

(e) Pearson's χ^2 goodness of fit test statistic is a realization from an approximate χ^2_9 distribution

(f) Pearson's χ^2 goodness of fit test statistic is 0

(g) Pearson's χ^2 goodness of fit test statistic is 18.05 rounded to the nearest two digits

(h) Pearson's χ^2 goodness of fit test statistic is 18.39 rounded to the nearest two digits

(i) Pearson's χ^2 goodness of fit test statistic is 213.01 rounded to the nearest two digits

(j) Pearson's χ^2 goodness of fit test statistic cannot be computed given the information available

(k) H_0 is rejected

(l) There is sufficient evidence to conclude this person is cheating on their tax return

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

Problem 3 [13min] (and 40min will have elapsed) Benford's law discovered in 1938 a discrete parameterless probability distribution on the first digits of numbers. It is also called the "first-digit law" and was discovered by examining real-world datasets. In real-world datasets, numbers beginning with a 1 are about twice as likely as numbers beginning with a 2. Numbers beginning with a 2 are about 50% more likely that numbers beginning with a 3, etc. This is especially true in domains such as accounting and taxes. This distribution is used frequently to prove that a dataset is fraudulent i.e. if someone is trying to cheat on their taxes by randomly generating numeric values, their values will not follow Benford's law since most crooks aren't too careful about learning probability theory. Below is the distribution's PMF and CDF. It's mean is 3.441.

x	1	2	3	4	5	6	7	8	9
$p(x)$	.301	.176	.125	.097	.079	.067	.058	.051	.046
$F(x)$	.301	.477	.602	.699	.778	.845	.903	.954	1.000

Consider a tax return with $n = 45$ numbers. We examine the first digit of the numbers and sort the data. It turns out the first digit is exactly uniformly distributed across all digits i.e. $\boldsymbol{x} =< 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, \ldots, 9, 9, 9, 9, 9 >$. This smells of fraud; we will investigate. To do so, we will employ the Kolmogorov-Smirnov (KS) test at $\alpha = 5\%$ (and assume it works for discrete rv's). The critical value you need for reference is: $F_K(1.36) = .95$ where K denotes the Kolmogorov distribution.

- [12 pt / 37 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) $H_0 : \theta = 3.441$ where θ denotes the mean of the DGP that generated $x_1, \ldots, x_{45}$

  (b) H_0 : the DGP that generated $x_1, \ldots, x_{45}$ is Benford's law

  (c) The KS test is an approximate test

  (d) The $\hat{\hat{D}}_n$ test statistic is 0

  (e) The $\hat{\hat{D}}_n$ test statistic is 0.19 rounded to the nearest two digits

  (f) The $\hat{\hat{D}}_n$ test statistic is 0.25 rounded to the nearest two digits

  (g) The $\hat{\hat{D}}_n$ test statistic is 0.27 rounded to the nearest two digits

  (h) The $\hat{\hat{D}}_n$ test statistic cannot be computed given the information available

  (i) If $\hat{\hat{D}}_n < 1.36$, this means the null hypothesis is retained

  (j) There is sufficient evidence to conclude this person is cheating on their tax return

  (k) the KS test statistic should be approximately equal to Pearson's χ^2 goodness of fit test statistic

  (l) Fisher's p-value in the KS test should be approximately equal to Fisher's p-value in Pearson's χ^2 goodness of fit test

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

Problem 4 [14min] (and 54min will have elapsed) Benford's law discovered in 1938 a discrete parameterless probability distribution on the first digits of numbers. It is also called the "first-digit law" and was discovered by examining real-world datasets. In real-world datasets, numbers beginning with a 1 are about twice as likely as numbers beginning with a 2. Numbers beginning with a 2 are about 50% more likely that numbers beginning with a 3, etc. This is especially true in domains such as accounting and taxes. This distribution is used frequently to prove that a dataset is fraudulent i.e. if someone is trying to cheat on their taxes by randomly generating numeric values, their values will not follow Benford's law since most crooks aren't too careful about learning probability theory. Below is the distribution's PMF and CDF. It's mean is 3.441.

x	1	2	3	4	5	6	7	8	9
$p(x)$	.301	.176	.125	.097	.079	.067	.058	.051	.046
$F(x)$	.301	.477	.602	.699	.778	.845	.903	.954	1.000

Consider a tax return with $n = 45$ numbers. We examine the first digit of the numbers and sort the data. It turns out the first digit is exactly uniformly distributed across all digits i.e. $\boldsymbol{x} = < 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, \ldots, 9, 9, 9, 9, 9 >$. This smells of fraud; we will investigate. To do so, we will employ the generalized likelihood ratio test at $\alpha = 5\%$. Critical values you may need to reference are: $F_{\chi^2_7}(14.1) = F_{\chi^2_8}(15.5) = F_{\chi^2_9}(16.9) = F_{\chi^2_{10}}(18.3) = .95$.

- [13 pt / 50 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) $H_0 : \theta = 3.441$ where θ denotes the mean of the DGP that generated $x_1, \ldots, x_{45}$

(b) H_0 : the DGP that generated $x_1, \ldots, x_{45}$ is Benford's law

(c) The $\hat{\hat{\Lambda}}$ test statistic is a realization from an approximate χ^2_9 distribution

(d) The generalized likelihood ratio test is an approximate test

(e) The numerator in the likelihood ratio is $1/9^{45}$

(f) The denominator in the likelihood ratio is $1/9^{45}$

(g) The likelihood that this data came from Benford's law is less than 1 in 1,000,000,000

(h) The $\hat{\hat{\Lambda}}$ test statistic is 0

(i) The $\hat{\hat{\Lambda}}$ test statistic is 2.12×10^{-47} rounded to the nearest two digits

(j) The $\hat{\hat{\Lambda}}$ test statistic is 8.59 rounded to the nearest two digits

(k) The $\hat{\hat{\Lambda}}$ test statistic cannot be computed given the information available

(l) If the likelihood ratio is calculated to be a value strictly greater than 1, the null hypothesis is rejected

(m) There is sufficient evidence to conclude this person is cheating on their tax return

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

Problem 5 [13min] (and 67min will have elapsed) Below is the PMF and log PMF of Benford's law distribution:

x	1	2	3	4	5	6	7	8	9
$p(x)$	.301	.176	.125	.097	.079	.067	.058	.051	.046
$\ln(p(x))$	-1.201	-1.737	-2.079	-2.333	-2.538	-2.703	-2.847	-2.976	-3.079

Consider a tax return with $n = 45$ numbers. We examine the first digit of the numbers and sort the data. It turns out the first digit is exactly uniformly distributed across all digits i.e. $\boldsymbol{x} =< 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, \ldots, 9, 9, 9, 9, 9 >$. This smells of fraud; we will investigate. This time, we will use model selection even though this is not, strictly speaking, a form of hypothesis testing. Consider two models:

I Uniform i.e. with likelihood $\left(\frac{1}{9}\right)^5 \left(\frac{1}{9}\right)^5 \cdot \ldots \cdot \left(\frac{1}{9}\right)^5$

II Benford i.e. with likelihood $(0.301)^5 (0.176)^5 \cdot \ldots \cdot (0.046)^5$

- [16 pt / 66 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) The log-likelihood in the first model is -98.88 rounded to the nearest two digits

  (b) The log-likelihood in the second model is -107.47 rounded to the nearest two digits

  (c) Since the log-likelihood in the second model is absolutely larger than the first model, the second model is more likely to be true if you assume one of the two models is true

  (d) The AIC for Model I is 197.75 rounded to the nearest two digits

  (e) The AIC for Model I is 199.75 rounded to the nearest two digits

  (f) Since Model I has the lower AIC, model I is the selected model according to the AIC model selection procedure

  (g) Assuming one of these two models is the true model, the probability that model I is true is 50%

  (h) Assuming one of these two models is the true model, the probability that model I is true is 99.98% rounded to the nearest two digits

  (i) For Model II, the AIC and the AICC metrics will be equivalent

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

Problem 6 [11min] (and 78min will have elapsed) Consider running $m > 1$ hypothesis tests. Following the notation from the lectures, below is a table that tabulates the random variables that model the possible events (denoted by uppercase letters) and the fixed constants (denoted by lowercase letters) in the course of these tests:

	Decision: Retain H_0	Decision: Reject H_0	total
H_0 true	U	V	m_0
H_a true	T	S	$m - m_0$
total	F	R	m

Assume each of the m tests are **independent from each other** and that the levels for each test are α.

- [13 pt / 79 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) $U \sim \text{Binomial}(m, \alpha)$

(b) $U \sim \text{Binomial}(m, 1 - \alpha)$

(c) For any $m - m_0$ alternative distributions, S can be modeled as a binomial

(d) T is the rv model which models the number of type II errors

(e) If the FWER is properly controlled, T goes to zero as $m \to \infty$

(f) If the FWER is properly controlled, the rejected tests's results are of higher practical significance then the rejected tests's results without FWER control (but still may not be practically significant)

(g) The familywise error rate (FWER) is equal to the probability that $R > 1$

(h) Using the Dunn-Sidak correction, the threshold for significance of each test would be less than α

(i) The Simes procedure has a higher $\mathbb{E}[R]$ than the Bonferroni procedure

(j) If you set $\alpha = 5\%/m$ then FWER $= 5\%$

(k) If you set $\alpha = 5\%/m$ then FWER $> 5\%$

(l) In order to compute the threshold of significance for Sime's FWER-controlling procedure, you need to first run all m tests and compute their m p-values

(m) For m_0 of the m tests, the p-values could be drawn as realizations from $\text{U}(0, 1)$

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

Problem 7 [9min] (and 87min will have elapsed) A not-so well-known test is a test for a DGP's "kurtosis". Kurtosis is generally speaking a measure of how fast the tails of the distribution approach zero. Pearson defined it to be $\kappa := \mathbb{E}\left[(X-\mu)^4/\sigma^4\right]$, the fourth standardized moment. A very interesting fact is that the normal distribution has $\kappa = 3$ regardless of its mean and variance. Thus, we define the "excess kurtosis" as the amount in excess over 3, i.e. $\kappa_0 := \kappa - 3$. An excess kurtosis of different than 0 means the tails of the distribution are thinner/fatter than the normal's tails. This is really important to test sometimes.

- [8 pt / 87 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) To test this we can use a one sample Wald test if we were given the DGP and can derive the MLE for κ_0

  (b) To test $H_a : \kappa_0 \neq 0$ we can use Pearson's χ^2 goodness of fit test

  (c) To test $H_a : \kappa_0 \neq 0$ we can use Kolmogorov-Smirnov's one-sample test

  (d) To test $H_a : \kappa_0 \neq 0$ we can use Fisher's two-sample permutation test

  (e) To test $H_a : \kappa_0 \neq 0$ we can use the multivariate delta method

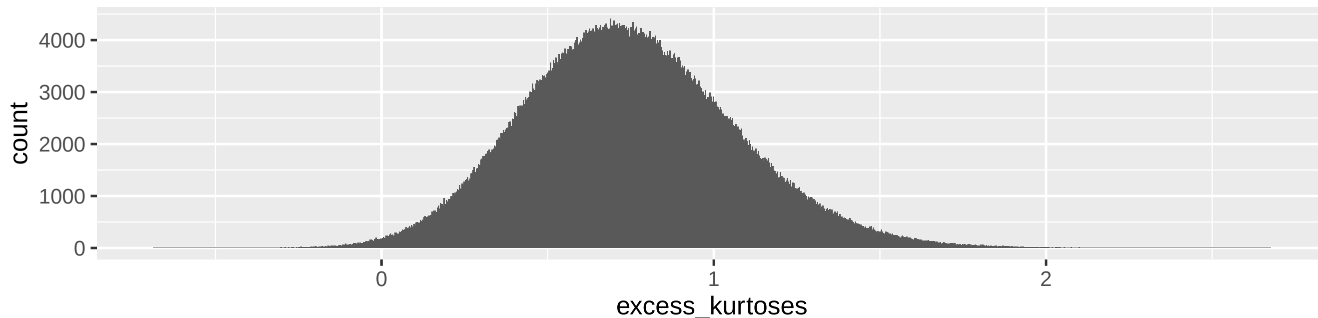
  (f) We can always estimate κ_0 using MM for any DGP whose κ_0 is finite

  (g) To test $H_a : \kappa_0 \neq 0$ we can use a one sample Wald test if we can derive the standard error for the MM κ_0 estimator

  (h) To test $H_a : \kappa_0 \neq 0$ we can use a bootstrap test where the MM excess kurtosis estimate is calculated for each bootstrap sample

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

Problem 8 [8min] (and 95min will have elapsed) A not-so well-known test is a test for a DGP's "kurtosis". Kurtosis is generally speaking a measure of how fast the tails of the distribution approach zero. Pearson defined it to be $\kappa := \mathbb{E}\left[(X-\mu)^4/\sigma^4\right]$, the fourth standardized moment. A very interesting fact is that the normal distribution has $\kappa = 3$ regardless of its mean and variance. Thus, we define the "excess kurtosis" as the amount in excess over 3, i.e. $\kappa_0 := \kappa - 3$. An excess kurtosis of greater than 0 means the tails of the distribution are fatter than the normal's tails. This is really important to test sometimes. Consider the following stock market data: centered daily percentage returns from the S&P 500 in the past year. We wish to test if the tails of this distribution are fatter than the normal distribution's tails, i.e. $H_a : \kappa_0 \neq 0$ at $\alpha = 5\%$ and to do so we'll use a bootstrap test. Here is a histogram of $B = 1,000,000$ MM estimates of κ_0. And $\hat{\hat{\kappa}}_0^{MM} = 0.748$



- [9 pt / 96 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

  (a) We do not really need to use the bootstrap test as the sampling distribution of $\hat{\kappa}_0^{MM}$ can be derived analytically

  (b) A 95% bootstrap confidence interval will be approximately between -1 and 3

  (c) A 95% bootstrap confidence interval will be approximately between 0.2 and 1.4

  (d) The null hypothesis is rejected

  (e) The data seems to suggest the normal model is not a good model for daily percentage returns from the S&P 500 in the past year

  (f) $B = 1,000,000$ seems to be sufficient to construct CI's and run hypothesis tests

  (g) $\hat{\kappa}_0^{MM}$ may be biased

  (h) If $\hat{\kappa}_0^{MM}$ were to be biased, the test is still valid regardless of the bias

  (i) If $\hat{\kappa}_0^{MM}$ were to be biased, the bias could be fixed by increasing n

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

Problem 9 [15min] (and 110min will have elapsed) Assume a DGP of $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(0, \theta)$. We wish to test $H_a : \theta \neq 1$ for the dataset 1.41, 1.44, 4.19, 5.12, 0.09, 8.62, 0.6, -8.88, 0.63, -8.57 at $\alpha = 5\%$ using the score test.

- [14 pt / 110 pts] Record the letter(s) of all the following that are **true** in general. At least one will be true.

(a) The score test requires you to compute $\hat{\hat{\theta}}^{MLE}$

(b) $\mathcal{L}(\theta; X_1, \ldots, X_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta} X_i^2\right)$

(c) $\ell(\theta; X_1, \ldots, X_n) = (2\pi\theta)^{-n/2} \ln\left(-\frac{1}{2\theta} X_i^2\right)$

(d) $\ell'(\theta; X_1, \ldots, X_n) = -\frac{n}{2\pi} - \frac{n}{2\theta} - \frac{1}{2\theta} \sum_{i=1}^{n} X_i^2$

(e) $\ell'(\theta; X_1, \ldots, X_n) = -\frac{n}{2\theta} + \frac{1}{2\theta^2} \sum_{i=1}^{n} X_i^2$

(f) $\ell''(\theta; X_1, \ldots, X_n) = \frac{n}{2\theta^2} - \frac{1}{\theta^3} \sum_{i=1}^{n} X_i^2$

(g) $I(\theta) = \frac{n}{2\theta^2}$

(h) $I(\theta) = -\frac{n}{2\theta^2}$

(i) The score statistic is $\left(-\frac{n}{2} + \frac{1}{2} \sum_{i=1}^{n} X_i^2\right) \frac{\sqrt{2}}{n}$

(j) The score statistic is 0

(k) The score statistic is 16.40 rounded to the two nearest digits

(l) The score statistic is 89.98 rounded to the two nearest digits

(m) The null hypothesis is rejected

(n) The p-value of the score test will be similar to the p-values of the likelihood ratio test and the wald test but may not be exactly equal

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