Econometrics 120A: Discussion Section

Midterm 2 Review

Linghui Wu & Lapo Bini

Department of Economics

Joint distributions

Joint distributions

Exercise 1: practice midterm Q2

$X \backslash Y$	0	1	2
0	0.05	0.1	0.2
1	0.1	0.1	0.1
2	0.2	0.1	0.05

(a) What is the marginal distribution of Y? What is the mean of Y?

(b) What is the conditional distribution of Y given $X = 1$? What is its mean?

(c) Are X and Y independent?

(d) Compute $\text{Cov}(X, Y)$.

(e) Compute $\text{Var}(X + Y)$.

Joint distributions

Exercise 1: practice midterm Q2

$X \backslash Y$	0	1	2
0	0.05	0.1	0.2
1	0.1	0.1	0.1
2	0.2	0.1	0.05

(a) What is the marginal distribution of Y? What is the mean of Y?

Joint distributions

Exercise 1: practice midterm Q2

$X \backslash Y$	0	1	2
0	0.05	0.1	0.2
1	0.1	0.1	0.1
2	0.2	0.1	0.05

(b) What is the conditional distribution of Y given $X = 1$? What is its mean?

Joint distributions

Exercise 1: practice midterm Q2

$X \setminus Y$	0	1	2
0	0.05	0.1	0.2
1	0.1	0.1	0.1
2	0.2	0.1	0.05

(c) Are X and Y independent?

Joint distributions

Exercise 1: practice midterm Q2

$X \backslash Y$	0	1	2
0	0.05	0.1	0.2
1	0.1	0.1	0.1
2	0.2	0.1	0.05

(d) Compute Cov(X, Y).

Joint distributions

Exercise 1: practice midterm Q2

$X \backslash Y$	0	1	2
0	0.05	0.1	0.2
1	0.1	0.1	0.1
2	0.2	0.1	0.05

(e) Compute Var$(X + Y)$.

Sampling and the Central Limit Theorem

Exercise 2: practice midterm Q3

A tour operator near a conference venue knows that there are 100 participants. They believe that if they price the tour at $20 per person, 80% of the participants will come. They get no revenue from participants who don't take the tour.

(a) What is the expected revenue?

(b) What is the (approximate) chance that, assuming their probabilities are correct, revenue falls below $1400?

(c) Suppose that they did see revenue below $1400. Does this indicate that their probability guess is too high?

(d) What is the chance that revenue is between $1500 and $1700?

(e) Suppose participants decide in groups of friends, so their decisions are not independent. Which part of your calculation in (b) is no longer correct? Is the true chance higher or lower than your answer?

Exercise 2: practice midterm Q3

A tour operator near a conference venue knows that there are 100 participants. They believe that if they price the tour at $20 per person, 80% of the participants will come. They get no revenue from participants who don't take the tour.

(a) What is the expected revenue?

Exercise 2: practice midterm Q3

A tour operator near a conference venue knows that there are 100 participants. They believe that if they price the tour at $20 per person, 80% of the participants will come. They get no revenue from participants who don't take the tour.

(b) What is the (approximate) chance that, assuming their probabilities are correct, revenue falls below $1400?

Exercise 2: practice midterm Q3

A tour operator near a conference venue knows that there are 100 participants. They believe that if they price the tour at $20 per person, 80% of the participants will come. They get no revenue from participants who don't take the tour.

(c) Suppose that they did see revenue below $1400. Does this indicate that their probability guess is too high?

Exercise 2: practice midterm Q3

A tour operator near a conference venue knows that there are 100 participants. They believe that if they price the tour at $20 per person, 80% of the participants will come. They get no revenue from participants who don't take the tour.

(d) What is the chance that revenue is between $1500 and $1700?

Exercise 2: practice midterm Q3

A tour operator near a conference venue knows that there are 100 participants. They believe that if they price the tour at $20 per person, 80% of the participants will come. They get no revenue from participants who don't take the tour.

(e) Suppose participants decide in groups of friends, so their decisions are not independent. Which part of your calculation in (b) is no longer correct? Is the true chance higher or lower than your answer?

Exercise 3: practice midterm Q4

x	-2	0	2
$P(X = x)$	0.2	0.6	0.2

We are interested in the random variable $R = \frac{1}{4}X^2$.

(a) What is the mean of R? What is its variance?

(b) We sample n observations from the distribution of R. We want the chance of a sample average greater than 0.5 to equal 0.05. What sample size n (rounded up to the next integer) is required?

(c) With $n = 100$ observations, what is the chance that the sample average of R is greater than 0.5?

(d) Compute $\text{Cov}(X, R)$. Are X and R independent?

Exercise 3: practice midterm Q4

x	-2	0	2
$P(X = x)$	0.2	0.6	0.2

We are interested in the random variable $R = \frac{1}{4}X^2$.

(a) What is the mean of R? What is its variance?

Exercise 3: practice midterm Q4

x	-2	0	2
$P(X = x)$	0.2	0.6	0.2

We are interested in the random variable $R = \frac{1}{4}X^2$.

(b) We sample n observations from the distribution of R. We want the chance of a sample average greater than 0.5 to equal 0.05. What sample size n (rounded up to the next integer) is required?

Exercise 3: practice midterm Q4

x	-2	0	2
$P(X = x)$	0.2	0.6	0.2

We are interested in the random variable $R = \frac{1}{4}X^2$.

(c) With $n = 100$ observations, what is the chance that the sample average of R is greater than 0.5?

Exercise 3: practice midterm Q4

x	-2	0	2
$P(X = x)$	0.2	0.6	0.2

We are interested in the random variable $R = \frac{1}{4}X^2$.

(d) Compute Cov(X, R). Are X and R independent?

Exercise 4: practice midterm Q5

The NSF polls US nationals every two years to evaluate scientific knowledge and beliefs. To the question "Does the Earth go around the Sun, or does the Sun go around the Earth?", 26 percent of those surveyed answered that the Sun goes around the Earth. Let $X_i = 1$ for this answer. The sample size was 2200.

(a) Most news agencies rounded the number to "one quarter", i.e. 25%. Suppose that 25% is the true proportion in the population. What is the chance that we would see a poll result of 26% or more?

(b) Is 25% a reasonable number for the true proportion in the population, given the survey results?

(c) If the true proportion is 25%, what is the chance of a poll result between 24% and 26%?

(d) How large would the sample need to be for a result of 26% or more to have only a 5% chance when the truth is 25%?

(e) Suppose the NSF had surveyed 2200 people from a single town. Which assumption of the CLT could fail, and what would it do to your answer in (a)?

Exercise 4: practice midterm Q5

The NSF polls US nationals every two years to evaluate scientific knowledge and beliefs. To the question "Does the Earth go around the Sun, or does the Sun go around the Earth?", 26 percent of those surveyed answered that the Sun goes around the Earth. Let $X_i = 1$ for this answer. The sample size was 2200.

(a) Most news agencies rounded the number to "one quarter", i.e. 25%. Suppose that 25% is the true proportion in the population. What is the chance that we would see a poll result of 26% or more?

Exercise 4: practice midterm Q5

The NSF polls US nationals every two years to evaluate scientific knowledge and beliefs. To the question "Does the Earth go around the Sun, or does the Sun go around the Earth?", 26 percent of those surveyed answered that the Sun goes around the Earth. Let $X_i = 1$ for this answer. The sample size was 2200.

(b) Is 25% a reasonable number for the true proportion in the population, given the survey results?

Exercise 4: practice midterm Q5

The NSF polls US nationals every two years to evaluate scientific knowledge and beliefs. To the question "Does the Earth go around the Sun, or does the Sun go around the Earth?", 26 percent of those surveyed answered that the Sun goes around the Earth. Let $X_i = 1$ for this answer. The sample size was 2200.

(c) If the true proportion is 25%, what is the chance of a poll result between 24% and 26%?

Exercise 4: practice midterm Q5

The NSF polls US nationals every two years to evaluate scientific knowledge and beliefs. To the question "Does the Earth go around the Sun, or does the Sun go around the Earth?", 26 percent of those surveyed answered that the Sun goes around the Earth. Let $X_i = 1$ for this answer. The sample size was 2200.

(d) How large would the sample need to be for a result of 26% or more to have only a 5% chance when the truth is 25%?

Exercise 4: practice midterm Q5

The NSF polls US nationals every two years to evaluate scientific knowledge and beliefs. To the question "Does the Earth go around the Sun, or does the Sun go around the Earth?", 26 percent of those surveyed answered that the Sun goes around the Earth. Let $X_i = 1$ for this answer. The sample size was 2200.

(e) Suppose the NSF had surveyed 2200 people from a single town. Which assumption of the CLT could fail, and what would it do to your answer in (a)?