Math Camp for Economists:

Expectation, population parameters, & sample statistics

Justin C. Wiltshire

Department of Economics
University of Victoria
Summer 2023

Overview of this lesson

In this lesson we're going to build up to several important concepts using what we've learned already and a few more basic concepts:

- Sequences
- Metric spaces and boundedness
- Limits
- Properties of the summation operator
- Arithmetic mean and the sample mean
- The expected value of a r.v. and properties of the expectation operator
- Variance, covariance, and correlation
- Statistics, estimators, and bias

Sequences

A "sequence" is an assignment of numbers to the natural numbers, with elements indexed from 1 to n, where n can be finite or infinite: $\{x_n\}_{n=1}^{c} = \{x_1, x_2, ..., x_n\}$

- A sequence of real numbers, $\{x_n\}$, is "increasing" ("decreasing") if $x_{n+1} \geq x_n$ ($x_{n+1} \leq x_n$) for all n
- If $\{x_n\}$ is a sequence of real numbers, then $\{x_n\}$ tends to infinity (written $\{x_n\}_{n \in \mathbb{N}}$ or $\{x_n\}_{n=1}^{\infty}$ or $x_n \to \infty$ or $\lim x_n = \infty$) if $\forall\ K \in \mathbb{R}\ \exists\ N(K)$ s.t. $n > N(K)$

Examples:

- $\{x_n\}_{n=1}^{6} = \{0, -1, 2, -3, 4, -5\}$
- $\{x_n\} = \{1, 2, 3, 4, ...\}$
- $\{x_n\}_{n \in \mathbb{N}} = \{1, -1, 2, -2, ...\}$
- $\{x_n\}_{n=1}^{\infty} = \{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, ...\}$

Metric spaces and boundedness

A "metric space" (X, d) generalizes the concept of distance between elements of a set

$\rightarrow$ X is a set and $d\colon X \times X \rightarrow \mathbb{R}_+$ is a function satisfying

(1) $d(x, y) \geq 0$, where $d(x, y) = 0 \iff x = y \; \forall \; x, y \in X$

(2) $d(x, y) = d(y, x) \; \forall \; x, y \in X$

(3) $d(x, z) \leq d(x, y) + d(y, z) \; \forall \; x, y, z \in X$ (triangle inequality)

In a metric space, (X, d), a subset $S \subseteq X$ is "bounded" if $\exists \; x \in X, \beta \in \mathbb{R}$ such that $\forall$ $s \in S$, $d(s, x) \leq \beta$

- A sequence $\{x_n\}$ is "bounded from below" if $\exists \; N_a$ such that $x_n \geq N_a \; \forall \; x_n$
- A sequence $\{x_n\}$ is "bounded from above" if $\exists \; N_b$ such that $x_n \leq N_b \; \forall \; x_n$

ε-balls

In a metric space, (X, d), define

- $\beta_\varepsilon(x) = \{y \in X \colon d(y, x) < \varepsilon\}$
  $= \text{an ``open ball'' with center } x \text{ and radius } \varepsilon$
- $\beta_\varepsilon[x] = \{y \in X \colon d(y, x) \leq \varepsilon\}$
  $= \text{a ``closed ball'' with center } x \text{ and radius } \varepsilon$

Limits, convergence, and divergence

A sequence of real numbers, $\{x_n\}$, has a "limit" point $L \in \mathbb{R}$ iff for each $\varepsilon > 0$, $\exists$ $N \in \mathbb{Z}^{++}$ such that if $n \geq N$ then $|x_n - L| < \varepsilon$

- A sequence $\{x_n\}$ "converges" to x ($x_n \to x$ or $\lim_{n\to\infty} x_n = x$) if it has a limit x
- If $\{x_n\}$ does not have a limit point, then we say it "diverges"
- If $x_n \to \infty$, then it diverges (or "is divergent")

Some practice

Exercise: Write down an example of a sequence that is strictly increasing and a sequence that is strictly decreasing. For each sequence:

(1) Does this sequence have a limit? If so, what is it?

(2) Write another example for which the opposite is true

(3) Explain to your classmates why these are the case

Summation operator

Suppose we have some sequence of numbers, $x_1, x_2, ..., x_n$

- We can write the sum of these numbers as $x_1 + x_2 + ... + x_n$
- The "summation operator" does this more compactly: $\sum\limits_{i=1}^{n} x_i = x_1 + x_2 + ... + x_n$

  → We will also often write it as $\sum_{i=1}^{n} x_i$ to save vertical space

Properties of the summation operator

There are many properties of summation operators. A few that are relevant for us:

i) For any constant c: $\sum_{i=1}^{n} c = nc$

ii) For any constant c: $\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i$

iii) For any constants c and d:
$\sum_{i=1}^{n}(c+d)x_i = \sum_{i=1}^{n} cx_i + \sum_{i=1}^{n} dx_i = c\sum_{i=1}^{n} x_i + d\sum_{i=1}^{n} x_i = (c+d)\sum_{i=1}^{n} x_i$

iv) For any constants c and d: $\sum_{i=1}^{n}(cx_i + dy_i) = c\sum_{i=1}^{n} x_i + d\sum_{i=1}^{n} y_i$

Exercise: Write down a concrete example of each of these properties

Properties the summation operator lacks

Summation operators lack many properties. A couple that are worth noting:

i) In general, $\sum_{i=1}^{n} \frac{x_i}{y_i} \neq \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i}$

→ This is not the case if $y_i = c \ \forall \ i$

· E.g. If $y_i = c \ \forall \ i$, then $\sum_{i=1}^{n} \frac{x_i}{y_i} = \sum_{i=1}^{n} \frac{x_i}{c} = \frac{1}{c} \sum_{i=1}^{n} x_i$

ii) In general, $\sum_{i=1}^{n} x_i^2 \neq (\sum_{i=1}^{n} x_i)^2$

Exercise: Write down a concrete example of how, in general, neither property holds

One more very useful property of summation operators

Now that we know, in general, $\sum_{i=1}^{n} x_i^2 \neq (\sum_{i=1}^{n} x_i)^2$, we note one more useful property of summation operators:

v) $\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n}(x_i - \bar{x})y_i$, where in general $\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$

Exercise:

(1) Show this is true (hint: what is $\sum_{i=1}^{n}(x_i - \bar{x})(-\bar{y})$?)

(2) What does this imply about $\sum_{i=1}^{n}(x_i - \bar{x})x_i$?

Arithmetic mean

The arithmetic mean, $\bar{x}$, is a measure of central tendency: it attempts to describe a data set by quantifying the center of the distribution such that the sum of deviations from $\bar{x}$ equals 0

Suppose we have a finite, discrete set of N data points $\{x_i\}_{i=1,...,N}$ where $x_i \neq x_j$ for ≥ 1 $i \neq j$, for $i, j \in N$

- Assume this set of data points has a central (i.e. "mean") value, $\bar{x}$, such that the sum of deviations from $\bar{x}$ equals 0
- Re-arrange the data so that $x_1, ..., x_M < \bar{x}$ and $x_{M+1}, ..., x_N \geq \bar{x}$, where $M < N$
- Then $(x_N - \bar{x}) + ... + (x_{M+1} - \bar{x}) - (\bar{x} - x_M) - ... - (\bar{x} - x_1) = 0$
- Rearranging this yields $x_1 + ... + X_N - N\bar{x} = \sum_{i=1}^{N} x_i - N\bar{x} = 0$

$$\implies \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

- This is true of any "sample", including a sample of observed values of r.v.s

See here for more discussion

Examples and exercises

If we have a sequence of N values, $x_1, ..., x_N$, then the mean is $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$

- Let $N = 3$, with $x_1 = 5$, $x_2 = 1$, $x_3 = 3$. Then $\bar{x} = \frac{1}{3}(5 + 1 + 3) = \frac{9}{3} = 3$
- Let $x_i = c$ some constant $\forall\ i$. Then $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} c = c\frac{N}{N} = c$

Exercise: Consider the sequence of numbers $\{x_1, x_2, x_3, x_4, x_5\} = \{2, 10, 15, 3, 5\}$

- Find $\bar{x}$, then show that $\sum_{i=1}^{N}(x_i - \bar{x}) = 0$

Another useful result

- We know $\sum_{i=1}^{N}(x_i - \bar{x}) = 0$, but we are often more interested in $\sum_{i=1}^{N}(x_i - \bar{x})^2$
- Unless $x_i = \bar{x} \; \forall \; i$, we know that $\sum_{i=1}^{N}(x_i - \bar{x})^2 > 0$
- Regardless, a useful result is that $\sum_{i=1}^{N}(x_i - \bar{x})^2 = \sum_{i=1}^{N} x_i^2 - N\bar{x}^2$
  - $\rightarrow$ More generally, $\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{N} x_i y_i - N\bar{x}\bar{y}$

Exercise:

(1) Why is it true that $\sum_{i=1}^{N}(x_i - \bar{x})^2 > 0$ unless $x_i = \bar{x} \; \forall \; i$? What if $x_i = \bar{x} \; \forall \; i$?

(2) Show that $\sum_{i=1}^{N}(x_i - \bar{x})^2 = \sum_{i=1}^{N} x_i^2 - N\bar{x}^2$

(3) Show that $\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{N} x_i y_i - N\bar{x}\bar{y}$

Population vs sample

- A “population” is the entire group in which you’re interested and about which you want to draw conclusions
- A “population parameter” summarizes an aspect of the population
- A “sample” is a subset of the population from which you collect data
  - $\rightarrow$ We’ll revisit this more in a little while

Expected value of a r.v.

Let X be a random variable, and let $E[\cdot]$ be the "expectation operator"

- The "expected value" of X is the weighted average of possible outcome values of X, $x_1, ..., x_K$
  - $\rightarrow$ Intuitively, the weights are the probability of each value occurring in the population
  - $\rightarrow$ That is, they are the population weights defined by the pmf or pdf associated with X, $f(x)$

  If X is a discrete r.v.

  $$\rightarrow \quad E[X] = x_1 f(x_1) + x_2 f(x_2) + ... + x_K f(x_K) = \sum_{j=1}^{K} x_j f(x_j)$$

  If X is a continuous r.v.

  $$\rightarrow \quad E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

We say $E[X] = \mu_X$ (or even just $E[X] = \mu$) is the "population mean" of X

$\rightarrow$ μ is a population parameter

Properties of $E[\cdot]$

If X is a r.v. with $E[X] = \mu_X$ and Y is a r.v. with $E[Y] = \mu_Y$,, and if $a, b \in \mathbb{R}$ are constants, then:

(1) $E[a] = a$

(2) $E[a + bX] = E[a] + bE[X] = a + b\mu$

(3) $E[b_1 X_1 + ... + b_K X_K] = b_1 E[X_1] + ... + b_K E[X_K]$

→ Expectations of weighted summations are summations of wtd expectations

→ Can compactly be expressed as $E\left[\sum_{i=1}^{K} a_i X_i\right] = \sum_{i=1}^{K} a_i E[X_i]$

(4) $E[E[X]] = E[X]$

→ $E[E[X]] = E[\mu_X] = \mu_X$

(5) $E[XY] = E[X]E[Y]$ if X and Y are independent

→ The converse is not true

Exercise: Write down a concrete example of each of these properties

More expected value exercises

Exercise: A r.v., Y, is the number of Tails from two flips of a fair coin

(1) What is the sample space?

(2) Write down $Y(\omega)$

(3) Write down the associated probability function, $f_Y(y)$

(4) What is $E[Y]$? Show your work

(5) Repeat (1) - (4) for $Z = Y^2$

Conditional expectation

Let X and Y be r.v.s, and A be an event in $\mathcal{F}$ with an associated non-zero probability

- The "conditional expectation" of a r.v. is its mean value in a population given some set of conditions occurs
- If X is a discrete r.v., then $E[X|A] = \sum_x xP(X = x|A)$
  $= \sum_x x\frac{P(\{X=x\}\cap A)}{P(A)}$
  $= 0$ if $P(A) = 0$
- If X and Y are discrete r.v.s, then $E[X|Y = y] = \sum_x xP(X = x|Y = y)$
  $= \sum_x x\frac{P(X=x,Y=y)}{P(Y=y)}$
  $\rightarrow$ $P(X = x, Y = y)$ is the "joint probability mass function" of X and Y
  $= 0$ if $P(Y = y) = 0$
- If X and Y are continuous r.v.s, then $E[X|Y = y] = \int_{-\infty}^{\infty} xf_{X|Y}(x|y)dx$
  $= \frac{1}{f_Y(y)}\int_{-\infty}^{\infty} xf_{X,Y}(x, y)dx$
  $\rightarrow$ $f_{X,Y}(x, y)$ is the "joint probability density function" of X and Y
  $\rightarrow$ Undefined if $f_Y(y) = 0$
- $E[XY|Y] = YE[X|Y]$
- $E[X] = E[E[X|Y]]$ by "the law of iterated expectations" (LIE)

The law of iterated expectations

Assume two r.v.s, X and Y, are discrete. Then

$$
\begin{aligned}
E[X] &= \sum_i x_i P(X = x_i) \\
&= \sum_i x_i \left(\sum_j P(X = x_i, Y = y_j) \right) \\
&= \sum_i x_i \left(\sum_j P(X = x_i | Y = y_j) P(Y = y_j) \right) \\
&= \sum_j \left(\sum_i x_i P(X = x_i | Y = y_j) \right) P(Y = y_j) \\
&= \sum_j E[X | Y = y] P(Y = y_j) \\
&= E\left[E[X | Y] \right]
\end{aligned}
$$

Variance of a r.v.

Let W be a r.v. with $E[W] = \mu_W$

- "Variance" is a measure of dispersion of a r.v. (how spread out it is from its mean)
  - $\rightarrow$ It is the squared deviation from the mean of a r.v.
  - $\rightarrow$ Squared because $E[W - E[W]] = E[W] - E[E[W]] = \mu_W - \mu_W = 0$
- $Var(W)$

$$\equiv V(W)$$
$$= Cov(W, W) \text{ is the "covariance" of } W \text{ with itself}$$
$$= E\left[(W - E[W])^2\right]$$
$$= E[(W - E[W])(W - E[W])]$$
$$= E[W^2] - 2E[E[W]W] + E[W]^2$$
$$= E[W^2] - 2E[W]E[W] + E[W]^2$$
$$= E[W^2] - E[W]^2$$
$$= \sum_{i=1}^{n}(w_i - \mu_W)^2 f(w_i) \text{ if W is discrete}$$
$$= \int_{-\infty}^{\infty}(w - \mu_W)^2 f(w)dw \text{ if W is continuous}$$
$$= \sigma_W^2 \text{ is the "population variance" of } W$$

  - $\rightarrow$ σ_W is the "population standard deviation" of W
  - $\rightarrow$ σ_W and σ_W^2 are "population parameters"

Properties of $V(\cdot)$

Let X and Y be r.v.s, and let $a, b \in \mathbb{R}, b \neq 0$ be constants

(1) $V(X) \geq 0$

(2) $V(a) = 0$

(3) $V(X) = 0 \iff \exists a\colon P(X = a) = 1$

(4) $V(a + bX) = b^2 V(X)$

(5) $V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$
$\quad = V(X) + V(Y)$ if X and Y are independent

(6) $V(aX - bY) = a^2 V(X) + b^2 V(Y) - 2abCov(X, Y)$

Exercise:

(1) Find $V(X - Y)$

(2) Show that $V(a + bX) = b^2 V(X)$

Variance exercises

Let $W(\Omega_W) = \{-1, 1\}$ and $Z(\Omega_Z) = \{-3, 2, 4\}$ be two discrete random variables, with associated pmfs

$$f_W(w) = \begin{cases} 1/3 & \textit{if } \; w = -1 \\ 2/3 & \textit{if } \; w = 1 \end{cases} \quad \text{and}$$

$$f_Z(z) = \begin{cases} 0.5 & \textit{if } \; z = -3 \\ 0.25 & \textit{if } \; z = 2 \\ 0.25 & \textit{if } \; z = 4 \end{cases}$$

Exercise:

(1) Find $E[W]$ and $E[Z]$

(2) Solve $= E\left[(W - E[W])^2\right]$ for $E[W^2] - E[W]^2$

(3) Find $V(W)$ and $V(Z)$

Standardized random variables

Let X be a r.v. with mean μ_X and variance $\sigma_X^2 > 0$.

- Then a new r.v., Z, is a "standardization" of X:

$$Z = \frac{X - \mu_X}{\sigma_X}$$

where $E[Z] = \mu_Z = 0$ and $V(Z) = \sigma_Z^2 = 1$

Exercise: Show that $E[Z] = 0$ and $V(Z) = 1$

Conditional variance

Let X and Y be r.v.s

- $V(X|Y) = E[X^2|Y] - E[X|Y]^2$
- $V(X) = V(E[Y|X]) + E[V(Y|X)]$

Covariance

Let X_1 and X_2 be two r.v.s

- "Covariance" is a measure of the joint variability of two r.v.s
- $Cov(X_1, X_2) = E\left[(X_1 - E[X_1])(X_2 - E[X_2])\right]$

  $= E[X_1 X_2] - E[X_1]E[X_2]$

  $\rightarrow \; = E[X_1^2] - E[X_1]^2 = V(X_1)$ if $X_2 = X_1$

  $\rightarrow \; = 0$ if X_1 and X_2 are independent

Let W and Z be the r.v.s defined earlier **Exercise**:

(1) Can you find $Cov(W, Z)$ without further information? What do you need to know? Show this?

(2) Find $Cov(W, W)$

Properties of $Cov(\cdot)$

Let X_1 and X_2 be two r.v.s, and let $a, b \in \mathbb{R}$ be constants. Define another r.v., $X_3 = aX_1 - bX_2$

- $Cov(X_1, a) = 0$
- $Cov(aX_1, bX_2) = abCov(X_1, X_2)$
- $Cov(X_1 + a, X_2 + b) = Cov(X_1, X_2)$

Exercise: Let $E[X_1] = \mu_1$, $E[X_2] = \mu_2$, $V(X_1) = \sigma_1^2$, and $V(X_2) = \sigma_2^2$, and let X_1, X_2 be independent

(1) Find $E[X_3]$

(2) Find $V(X_3)$

(3) Find $Cov(X_1, X_3)$

(4) Find $V(aX_1 - bX_3)$

Correlation

The magnitude of the covariance of two r.v.s does not have a straightforward interpretation. For this, "correlation"—a measure of a relationship between two things—can be more useful

Let X and Y be r.v.s, and define $W = \frac{X - E[X]}{\sqrt{V(X)}}$ and $Z = \frac{Y - E[Y]}{\sqrt{V(Z)}}$

- $Corr(X, Y) = Cov(W, Z) = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}} = \rho_{XY}$ if $\sqrt{V(X)V(Y)} \neq 0$
  - $\rightarrow$ $\rho_{XY} \in [-1, 1]$
  - $\rightarrow$ If $\rho_{XY} = 0$, we say X and Y are "uncorrelated"
    - $\not\Rightarrow$ X, Y independent!
  - $\rightarrow$ However, X, Y independent $\Rightarrow$ X, Y uncorrelated

Exercise: X_1, X_2, and X_3 were defined earlier. Find $\rho_{X_1 X_3}$ if X_1 and X_2 are independent

Random samples and statistics

- A sequence of r.v.s, $X_1, ..., X_n$ are a "random sample" of size n from a population if they are "independent and identically distributed" r.v.s
  - $\rightarrow$ Here, "independent" means each X_i is an independent event
  - $\rightarrow$ "identically distributed" means the X_i share a common probability space $(\Omega, \mathcal{F}, P)$ and distribution
- Let $X_1, ..., X_n$ be a random sample from population, and let $T(x_1, ..., x_n)$ be a real-valued function with a domain that includes the sample space of $(X_1, ..., X_n)$
  - $\rightarrow$ Then the r.v. $Y = T(X_1, ..., X_n)$ is called a "statistic"
  - $\rightarrow$ The prob dist'n of a statistic $Y = T(X)$ is the "sampling distribution" of Y
- An "estimator" is a statistic, $\hat{\theta}$, used to infer an unknown population parameter, θ, in a statistical model (e.g. μ or σ)

Some common statistics

Let $X_1, \ldots, X_n$ be a "random sample" of size n from a population

- The "sample mean" is the arithmetic mean of the values in a random sample, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - → $\bar{x}$ denotes the "observed value" of $\bar{X}$ in any random sample
- The "sample variance" measures the dispersion of the sample data from the sample mean. We will define $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
  - → s^2 denotes the "observed value" of S^2 in any random sample
  - → We will see shortly why we use $n-1$ in the denominator instead of n

Estimators and bias

An estimator, $\hat{\theta}$, is a statistic used to infer an unknown population parameter, θ

- "Bias" in an estimator is the difference between the expected value of an estimator and the true value of the population parameter being estimated
  $\rightarrow \; Bias(\hat{\theta}, \theta) = E\left[\hat{\theta}\right] - \theta$

Example: Let $X_1, ..., X_n$ be *iid* r.v.s with mean μ and variance σ^2, and define $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

- Is $\bar{X}$ a biased estimator of μ?

$$\begin{aligned}
\mathrm{E}\left[\bar{X}\right] &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} E[X_i] \\
&= \frac{n}{n}\mu \\
&= \mu \\
&\iff Bias(\bar{X}, \mu) = \mathrm{E}\left[\bar{X}\right] - \mu = \mu - \mu = 0
\end{aligned}$$

So $\bar{X}$ is an "unbiased estimator" of μ

Lab exercise 1

Exercise:

A r.v., Y, is the face value of one roll of a fair four-sided die

(1) What is the sample space?

(2) Write down $Y(\omega)$

(3) Write down the associated probability function, $f_Y(y)$

(4) What is $E[Y]$? Show your work

(5) Repeat (1) - (4) for $Z = Y^2$

Lab exercise 2

Exercise:

Let a r.v. $Y(\omega) = \begin{cases} -2 & \text{if } \omega = H \\ 0 & \text{if } \omega = T \end{cases}$, is the value of one flip of a two-sided coin

(1) Write down the probability function, $f_Y(y)$, associated with $Y(\omega)$

(2) What is $E[Y]$? Show your work

(3) Repeat (1) - (4) for $Z = Y^2$

Lab exercise 3

Provide a proof for the law of iterated expectations i.e. $E[Y] = E[E[Y|X]]$

Lab exercise 4 (repeat of in-lesson exercises)

Exercise: Write down an example of a sequence that is strictly increasing and a sequence that is strictly decreasing. For each sequence:

(1) Does this sequence have a limit? If so, what is it? Write another example for which the opposite is true

(2) Show that $\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n}(x_i - \bar{x})y_i$

(3) Consider the sequence of numbers $\{x_1, x_2, x_3, x_4, x_5\} = \{2, 10, 15, 3, 5\}$

- Find $\bar{x}$, then show that $\sum_{i=1}^{N}(x_i - \bar{x}) = 0$

(4) Why is it true that $\sum_{i=1}^{N}(x_i - \bar{x})^2 > 0$ unless $x_i = \bar{x} \; \forall \; i$? What if $x_i = \bar{x} \; \forall \; i$?

(5) Show that $\sum_{i=1}^{N}(x_i - \bar{x})^2 = \sum_{i=1}^{N} x_i^2 - N\bar{x}^2$

(6) Show that $\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{N} x_i y_i - N\bar{x}\bar{y}$

Lab exercise 5

Exercise:

A r.v., Y, is the face value of one roll of a fair six-sided die

(1) What is the sample space?

(2) Write down $Y(\omega)$

(3) Write down the associated probability function, $f_Y(y)$

(4) What is $E[Y]$? What is $V(Y)$? Show your work

(5) Repeat (1) - (4) for $Z = Y^2$

(6) What is $Cov(Y, Z)$?

(7) What is $Corr(Y, Z)$?

(8) Are Y and Z independent?

Lab exercise 6

Exercise:

Let a r.v. $Y(\omega) = \begin{cases} -1 & \text{if } \omega = H \\ 1 & \text{if } \omega = T \end{cases}$, is the value of one flip of a two-sided coin

(1) Write down the probability function, $f_Y(y)$, associated with $Y(\omega)$

(2) What is $E[Y]$? What is $V(Y)$? Show your work

(3) Repeat (1) - (4) for $Z = Y^2$

(4) What is $Cov(Y, Z)$?

(5) What is $Corr(Y, Z)$?

(6) Are Y and Z independent?

Lab exercise 7

Exercise:

Let $X_1, \ldots, X_n$ be *iid* r.v.s with mean μ and variance σ^2, and define $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\tilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$. We showed, earlier, that $E\left[\bar{X}\right] = \mu$, thus $\bar{X}$ is an unbiased estimator of μ

(1) Define a statistic $\tilde{X} = \frac{1}{n-1}\sum_{i=1}^{n} X_i$. Is $\tilde{X}$ also an unbiased estimator of μ? Why?

(2) What is the variance of $\bar{X}$?

(3) Is $\tilde{S}^2$ a biased estimator of σ^2? It may be helpful to recall that:

   a) Adding and subtracting $\bar{X}$ from $X_i - \mu$ yields $X_i - \mu = (X_i - \bar{X}) + (\bar{X} - \mu)$
   b) $E[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2] = V(X_i)$
   c) $E[\frac{1}{n}\sum_{i=1}^{n}(\bar{X} - \mu)^2] = V(\bar{X})$, which you found in (2)

(4) Given your answer to (3), define an unbiased estimator for σ^2. Use your answer to (3) to show that this estimator is indeed unbiased

Lab exercise 8 (Repeat from lesson)

Let X and Y be two r.v.s, and let $a, b \in \mathbb{R}$ be constants

(1) Find $V(X - Y)$

(2) Show that $V(a + bX) = b^2 V(X)$

(3) If a r.v. X has mean μ_X and variance σ_X^2, define a new r.v. Z as a standardization of X. Show that Z, the standardization of X, has mean 0 and variance 1

Lab exercise 9 (Repeat from lesson)

Let X_1 and X_2 be two independent r.v.s, with $E[X_1] = \mu_1$, $E[X_2] = \mu_2$, $V(X_1) = \sigma_1^2$, and $V(X_2) = \sigma_2^2$, and let $a, b \in \mathbb{R}$ be constants. Define another r.v., $X_3 = aX_1 - bX_2$.

(1) Find $E[X_3]$

(2) Find $V(X_3)$

(3) Find $Cov(X_1, X_3)$

(4) Find $V(aX_1 - bX_3)$

(5) Find $\rho_{X_1 X_3}$

Lab exercise 10 (Repeat from lesson)

Let $W(\Omega_W) = \{-1, 1\}$ and $Z(\Omega_Z) = \{-3, 2, 4\}$ be two discrete random variables, with associated pmfs

$$f_W(w) = \begin{cases} 1/3 & \textit{if } w = -1 \\ 2/3 & \textit{if } w = 1 \end{cases} \quad \text{and}$$

$$f_Z(z) = \begin{cases} 0.5 & \textit{if } z = -3 \\ 0.25 & \textit{if } z = 2 \\ 0.25 & \textit{if } z = 4 \end{cases}$$

(1) Find $E[W]$ and $E[Z]$

(2) Solve $= E\left[(W - E[W])^2\right]$ for $E[W^2] - E[W]^2$

(3) Find $V(W)$ and $V(Z)$

(4) Can you find $Cov(W, Z)$ without further information? What do you need to know? Show this?

(5) Find $Cov(W, W)$

References

This drew notes from:

- *Causal Inference: The Mixtape* (Cunningham, 2021)
- Wikipedia
- A stackexchange thread
- Brown University Math Camp notes
- UC Berkeley Math Camp notes
- Iowa State Basic Statistics notes