STAB57: An Introduction to Statistics

Shahriar Shams

Week 1 (Introduction and Review of STAB52)

Winter 2023

Acknowledgement

I am grateful to these people who helped me create the content of this course...

- **Dr. Mahinda Samarakoon**, Professor, Department of Computer and Mathematical Sciences, University of Toronto Scarborough.
- **Dr. Jabed Tomal**, Assistant Professor, Department of Mathematics and Statistics, Thompson River University.
- **Dr. Jamie Stafford**, Professor, Department of Statistical Sciences, University of Toronto.
- **Alex Stringer**, Assistant Professor, Department of Statistics and Actuarial Science, University of Waterloo.

Review: Probability

- The **probability measure** P for each event A defined on sample space Ω satisfies the following properties:
  - $P(A)$ is non-negative and $0 \leq P(A) \leq 1$
  - $P(A) = 0$ when A is empty
  - $P(A) = 1$ when A is the entire sample space Ω
  - P is (*countably*) additive.

X is the outcome of rolling a fair dice.
What is the probability that it's an even number?

$\Omega = \{1,2,3,4,5,6\}$, $A=\{2,4,6\}$
$\implies P(A) = 3/6 = 1/2$

Review: Expectation

- **Expected value/ mean/ average** of random variable (X) is defined as
  - $E[X] = \int_{-\infty}^{\infty} x f(x) dx$ when X is continuous or
  - $E[X] = \sum_i x_i P[X = x_i]$ when X is discrete

X is the outcome of rolling a fair dice.
What is the expected value of X?

$$E[X] = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6} = \frac{1+2+...+6}{6} = 3.5$$

- Expectation is a **linear operator**
  - Let X and Y are two random variables and a, b and c are few constants. Then
  - $E[aX + bY + c] = aE[X] + bE[Y] + c$

Indicator Function

- If A is any event, we can define the **indicator function of A, written I_A**, to be the random variable for all $s \in \Omega$

$$I_A(s) = \begin{cases} 1, \text{ if s} \in A \\ 0, \text{ if s} \notin A \end{cases}$$

Probability expressed as the expectation of Indicator function

Using the same example as before: We are rolling a dice and $A = \{2, 4, 6\}$

Random variable X	1	2	3	4	5	6
I_A	0	1	0	1	0	1

$E[I_A] = \frac{1}{6}(0 + 1 + 0 + 1 + 0 + 1) = \frac{3}{6} = \frac{1}{2} = P[A]$

Review: LLN

- **Law of Large Number (LLN)**
  - Let $X_1, X_2, ..., X_i$ be a sequence of independent random variables with $E[X_i] = \mu$.
  - Let $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$
  - Then $\bar{X}_n \xrightarrow{P} \mu$ as $n \to \infty$
  - **In naive words:** sample mean approaches the population mean as the sample size increases.

We are rolling a fair dice repeatedly and calculating the mean

Sample size (n)	Observations	Sample mean ($\bar{X}_n$)
3	3,4,1	8/3=2.67
5	3,4,1,6,5	$19/5 = 3.8$
...	...	...
800	3,4,1,...,2,5	3.49

NOTE: population average = 3.5

LLN in graph

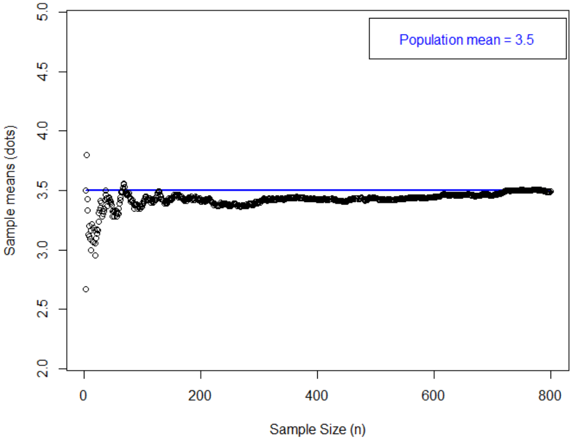

Figure: Trace of sample mean from repeatedly rolling a fair dice

Review: Central Limit Theorem (CLT)

- Suppose $X_1, X_2, ...$ is an i.i.d. sequence of random variables each having finite mean μ and finite variance σ^2
- Let $\bar{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ i.e. sample mean
- Then according to the Central Limit Theorem as $n \to \infty$,

$$\bar{X}_n \xrightarrow{D} N(\mu, \frac{\sigma^2}{n})$$

or

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0, 1)$$

- **In naive words:** A random variable (X) can follow *some distribution* with mean μ and variance σ^2. If we pick a *fixed number of samples (n)* and *calculate the sample mean repeatedly*, then those sample means will have a **Normal distribution** with mean μ and variance σ^2/n

Review: Linear Combination of Normal variables

- Let $X_i \sim N(\mu_i, \sigma_i^2)$ where $i = 1, 2, ...n$ and X_i's are independent.
- Let Y be a linear combination of all the X_i's with

$$Y = a_1 X_1 + a_2 X_2 + ...a_n X_n + b = \sum_{i=1}^{n} a_i X_i + b$$

  where $a_1, a_2, ..., a_n, b$ are constants

- Then,

$$Y \sim N\left(\sum_{i=1}^{n} a_i \mu_i + b\,,\ \sum_{i=1}^{n} a_i^2 \sigma_i^2\right)$$

Example

Let, $X_1 \sim N(10, 2)$ and $X_2 \sim N(20, 3)$ and $Y = 0.4X_1 + 0.6X_2$
Then $Y \sim N(\ ,\ )$ with mean$=0.4 * 10 + 0.6 * 20 = 16$ and
variance $= (0.4)^2 * 2 + (0.6^2) * 3 = 1.4$

Review: Some common distributions

Distribution	**pdf or pmf**	**mean**	**variance**	**MGF**
Bernoulli(θ)	$\theta^x(1-\theta)^{1-x}$	θ	$\theta(1-\theta)$	$(1-\theta)+\theta e^t$
Binomial (m,θ)	$\binom{m}{x}\theta^x(1-\theta)^{m-x}$	$m\theta$	$m\theta(1-\theta)$	$[(1-\theta)+\theta e^t]^m$
Poisson(λ)	$\frac{e^{-\lambda}\lambda^x}{x!}$	λ	λ	$exp[\lambda(e^t-1)]$
Uniform$[a,b]$	$1/(b-a)$	$(a+b)/2$	$(b-a)^2/12$	$\begin{cases} (e^{tb}-e^{ta})/t(b-a) & ,t \neq 0 \\ 1 & ,t=0 \end{cases}$
Normal (μ,σ^2)	$(2\pi\sigma^2)^{-1/2}exp[-\frac{1}{2\sigma^2}(x-\mu)^2]$	μ	σ^2	$exp[\mu t+\sigma^2t^2/2]$
Exponential(β)	$\beta e^{-\beta x}$	$1/\beta$	$1/\beta^2$	$(1-t/\beta)^{-1}$
Exponential(θ)	$\frac{1}{\theta}e^{-\frac{x}{\theta}}$	θ	θ^2	$(1-t\theta)^{-1}$
Gamma(α,β)	$\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$	α/β	α/β^2	$(1-t/\beta)^{-\alpha}\,,t<\beta$
Gamma(α,θ)	$\frac{1}{\Gamma(\alpha)\theta^\alpha}x^{\alpha-1}e^{-\frac{x}{\theta}}$	$\alpha\theta$	$\alpha\theta^2$	$(1-\theta t)^{-\alpha}\,,t<1/\theta$
Beta(a,b)	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$	$a/(a+b)$	$\frac{ab}{(a+b)^2(a+b+1)}$	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$

Review: Z and χ^2 distribution

- **Standard Normal/ $N(0,1)$ / Z distribution**
  - If $X \sim N(\mu, \sigma^2)$ then $\frac{X-\mu}{\sigma} \sim N(0,1)$
  - $Z = \frac{X-\mu}{\sigma}$
- **χ^2 distribution**
  - Let $U = Z^2$ where Z is a Standard Normal variable
  - $U \sim \chi^2$ distribution with 1 degrees of freedom. (written as $\chi^2_{(1)}$)
  - Additive property: If $X \sim \chi^2_{(m)}$ and $Y \sim \chi^2_{(n)}$ then $X + Y \sim \chi^2_{(m+n)}$
  - If $X \sim \chi^2_{(m)}$ then $E[X] = m$

Review: t and F distribution

- **t distribution**
  - Let Z and U are two independent variables
  - where $Z \sim N(0,1)$ and $U \sim \chi^2_{(m)}$
  - $\frac{Z}{\sqrt{U/m}} \sim t$-distribution with m degrees of freedom. (written as $t_{(m)}$)
- **F distribution**
  - Let X and Y are two independent variables
  - where $X \sim \chi^2_{(m)}$ and $Y \sim \chi^2_{(n)}$
  - Then $\frac{X/m}{Y/n} \sim F$ distribution with degrees of freedom (m, n)

Homework (Non-credit)

Evans and Rosenthal

Exercise: 3.4.21, 3.4.23, 4.6.1 - 4.6.10

Rice

Chapter 6, Exercise: 3, 5, 6