Homework 7

7.1 Revisiting Super Bowl Coin Flips

Recall the Super Bowl coin flip problem. Assume the random variable X is defined as:

$$X = \begin{cases} 0 & NFC\ wins\ coin\ flip \\ 1 & AFC\ wins\ coin\ flip \end{cases}$$

(a) If the coin were fair (50/50 chance of winning a flip), what is the theoretical:

i) Probability mass function p(x) (Hint: don't over-think this)

Solution

PMF for all the values p(x) = 1/2

ii) Population mean $\mu_0 = \sum_{I=0}^{1} p(i).i$

Solution

Mean = 0.5

iii) Population standard deviation σ_0 or variance $\sigma_0^2 = \sum_{i=0}^{1} p(i) \cdot (i - \mu_0)^2$

Solution

σ_0^2 = 0.5 x $(1 – 0.5)^2$ + 0.5 x $(0 – 0.5)^2$

= 0.25

σ_0 = 0.5

(b) Compute the following for N = 52 observed values of X in superbowl.csv:

i) Sample mean $\overline{x}$

Solution

$\overline{x}$ = 0.3269

ii) Sample standard deviation s_x or variance s_x^2

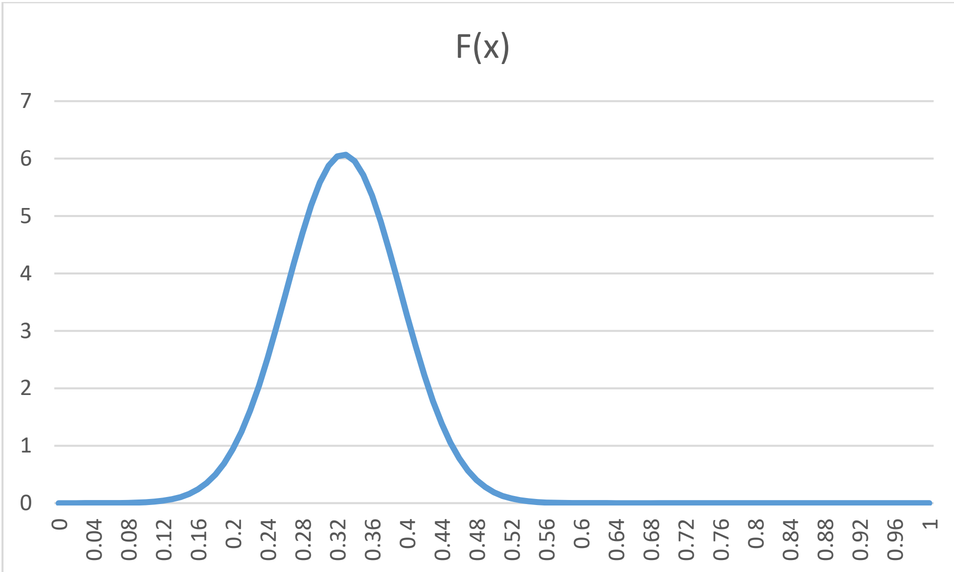
Solution

S_x = 0.473

s_x^2 = 0.2243

(c) Create a plot of the PDF $f(\bar{x})$ for values $0 \leq \bar{x} \leq 1$ using the Central Limit Theorem to model the distribution of sample means for N = 52 trials.

Solution



(d) Perform a hypothesis test for the following:

$$H_0: \mu_x = \mu_0$$
$$H_a: \mu_x \neq \mu_0$$

Report the p-value and determine whether H_0 can be rejected at $\alpha = 0.05$.

Solution

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{N}}$$

∴ z = -2.49

$$p = 2 \times (1 - F_{norm}(|z|)$$
$$\therefore p = 0.012$$

We cannot accept H_0

(e) What can you conclude about the validity of the Superbowl coin flip?

Solution:

The value of z and p are not normal. Therefore we can assume that the coin flip is one-sided.

7.2 GRE Tutoring Service

A $1799 tutoring service advertises a significant increase in verbal reasoning GRE score. The attached file gre.csv contains a set of N = 100 samples of pre- and post-test scores for participating students.

Sample Space

	pre	post
N	100	100
mean	149.08	150.9
std div (Sample)	8.829107	10.30593

(a) Assuming the pre- and post-test data are not related (i.e. randomly ordered), perform a hypothesis test for the following:

$$H_0: \mu_{pre} = \mu_{post}$$
$$H_a: \mu_{pre} < \mu_{post}$$

Report the p-value and determine whether H_0 can be rejected at α = 0.05

Solution:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - \delta}{\sqrt{\frac{\sigma_1^2}{N_1} - \frac{\sigma_2^2}{N_2}}}$$

$$z = \frac{(150.9 - 149.08) - 0}{\sqrt{\frac{10.31^2}{100} - \frac{8.83^2}{100}}}$$

$$\therefore z = 1.34$$

P = F_{norm} (|z|)

=0.089

When α = 0.05 the z bounds are between [-1.96, 1.96] so we can assume the hypothesis to be correct.

(b) Do the results in (a) support the tutoring service's advertising claim?

Solution:

Yes, the advertisement claim and the dataset support each other.

(c) Assuming the pre- and post-test data are related (i.e. paired from the same student), perform a hypothesis test for the following:

$$H_0: \mu_{pre} = \mu_{post}$$
$$H_a: \mu_{pre} < \mu_{post}$$

Report the p-value and determine whether H_0 can be rejected at $\alpha = 0.05$

Solution

p-value = 0.0034 {using t test}

The p-value is too small to be considered for the hypothesis to be true.

(d) Do the results in (c) support the tutoring service's advertising claim?

Solution:

The claim seems to be false. There is no significant difference in the dataset.

(e) Are the results in (c) practically significant? Would you buy the service? Why?

Solution:

There is very less significant increase in marks & are therefore not practically significant. I wouldn't buy these services because there are very few people who benefit from these services.