GRADUATION THESIS POSTER EVENT

ON THE DISTRIBUTION OF SAMPLING STATISTICS

MATHEMATICS, YILDIZ TECHNICAL UNIVERSITY, ISTANBUL

MELİS PATAR 18025015

Advisor: Asst. Prof. ÖZLEM BAKŞİ

2022-2023 Fall Term

The Sample Mean and Variance

Consider a population of elements, each of which has a numerical value attached to it. For instance, the population might consist of the adults of a specified community and the value attached to each adult might be his or her annual income, or height, or age, and so on. The sample mean is defined by

$$\overline{X} = \frac{X_1+...+X_n}{n}$$

EXAMPLE

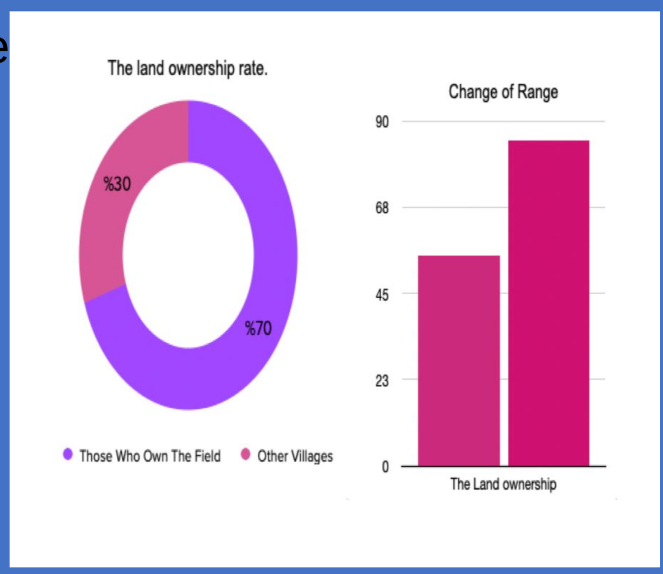
A mature sequoia tree has an average height of 220 feet with a standard deviation of 25. The table shows a random sampling of 30 trees whose heights were measured. How does the mean for this data set compare with the population mean?

$$\overline{X} = \frac{X_1+...+X_n}{n} = \frac{6,84464}{30} \simeq 228.15$$

231.27	230.71	229.52
281.31	296.80	219.18
228.11	197.82	231.27
214.89	271.63	239.01
208.69	227.42	234.17
231.92	178.51	230.81
214.13	230.96	206.74
223.77	234.07	224.30
195.28	259.81	236.35
209.61	202.38	224.22

Approximate Distribution of the Sample Mean Example

It is known that 70% of the people living in a village own their own fields. (a) What is the probability that more than 75% of the inhabitants of this village have their own field?
(b) The ratio of those living in this village who own their own fields is between 55% and 85%. What is the probability that it will happen?



(a) Firstly, the sampling distribution of the random variable P should be determined.

$$E(P) = p = 0.70$$

$$Var(P) = \frac{p(1-p)}{n} = \frac{0.70\times0.30}{100} = 0.0021$$

and it can be said that $P \sim N(0.70, 0.0021)$ from the information given above. Accordingly, the requested probability is found in the standard normal distribution table as

$$P(P \geq 0.75) = P(\frac{P-0.70}{\sqrt{0.0021}} \geq \frac{0.75-0.70}{\sqrt{0.0021}})$$

$$= P(Z \geq \frac{0.75-0.70}{\sqrt{0.0021}}) = 1.0910$$

$$P(0.55 \leq P \leq 0.85) = P(\frac{0.55-0.70}{\sqrt{0.0021}} \leq \frac{P-0.70}{\sqrt{0.0021}} \leq \frac{0.75-0.70}{\sqrt{0.0021}})$$

$$=P(\frac{0.55-0.70}{\sqrt{0.0021}} \leq Z \leq \frac{0.75-0.70}{\sqrt{0.0021}}) = P(-3.2732 \leq Z \leq 1.910).$$

is calculated as.⊕

The Central Limit Theorem

Let $X_1, X_2,..., X_n$ be independent random variables with the same distribution with expected value μ variance σ^2.
Here, we will introduce and moment generating functions (MGFs). Moment generating functions are useful for several reasons, one of which is their application to analysis of sums of random variables. Before discussing MGFs, let's define moments.The *n*th moment of a random variable X defined to be $E[X^n]$. The *n*th central moment of X is defined to be $E[((X - E[X])^n)]$.
If the moment generating function of random variables is in the vicinity of the zero point, if $n \to \infty$, it is

$$\frac{\sqrt{n}(\overline{X}_n-\mu)}{\sigma} \to N(0,1).$$

EXAMPLE

Let's calculate the probability of getting heads at least 60 times when a coin is tossed 100 times.This probability can be easily calculated from the Binomial Expansion. Since , the probability sought is $P(x \geq 60)$ and binomial random variable having parameters (n, p) is given by ,

$$p(i) = \binom{n}{i} p^i (1-p)^{n-i}, \qquad i = 0, 1, \ldots, n$$

This probability can be calculated as

$$P(x \geq 60) = \sum_{x=60}^{100} P(X = x) = \sum_{x=60}^{100} (\frac{100}{x})(\frac{1}{2})^x(\frac{1}{2})^{100-x} = (\frac{1}{2})^{100} \sum_{x=60}^{100} \frac{100}{x} = 0.02844$$

is calculated.. The same probability can be approximated by this theorem (The Central Limit Theorem). Let X_i be the number of heads in each toss of the coin. Where $X := X_1 + X_2 +... + X_{100}$ is the total number of heads in 100 flips of the coin. Also we know that from Bernoulli Random Variable $X - P(x \geq 60)$. Trials are independent of each other $E(X) = 100(1/2) = 50$ are $Var(X) = 100(1/2)(1/2) = 25$.

Given that 100 trials are large enough for this example, the probability $P(x \geq 60)$ is approximated to the

$$P(x \geq 60) = \sum_{x=60}^{100} (X_i \geq 60) = P\left(\frac{\sum_{x=60}^{100}(X_i - E\left[\sum_{x=60}^{100} X_i\right]}{\sqrt{Var\left[\sum_{x=60}^{100} X_i\right]}} \geq \frac{60-50}{5}\right) \cong P(Z \geq 2) = 0.0228$$

shape according to the theorem. The difference is very small. The larger the sample size n, the smaller the difference.⊕

SAMPLING DISTRIBUTIONS FROM A NORMAL POPULATION

In the 2012 Presidential Election, President Obama received 52% of the vote in Pennsylvania. On the day of the election the outcome in Pennsylvania was important to the national election outcome so before all of the votes were counted, several pollsters conducted "exit polls" to gauge how the vote turned out and the reasons why people voted as they did. Suppose you conduct an exit poll of 1000 Pennsylvania voters leaving their precinct voting stations or after they had voted by mail. What is the probability that a majority of your sample did not vote for President Obama?

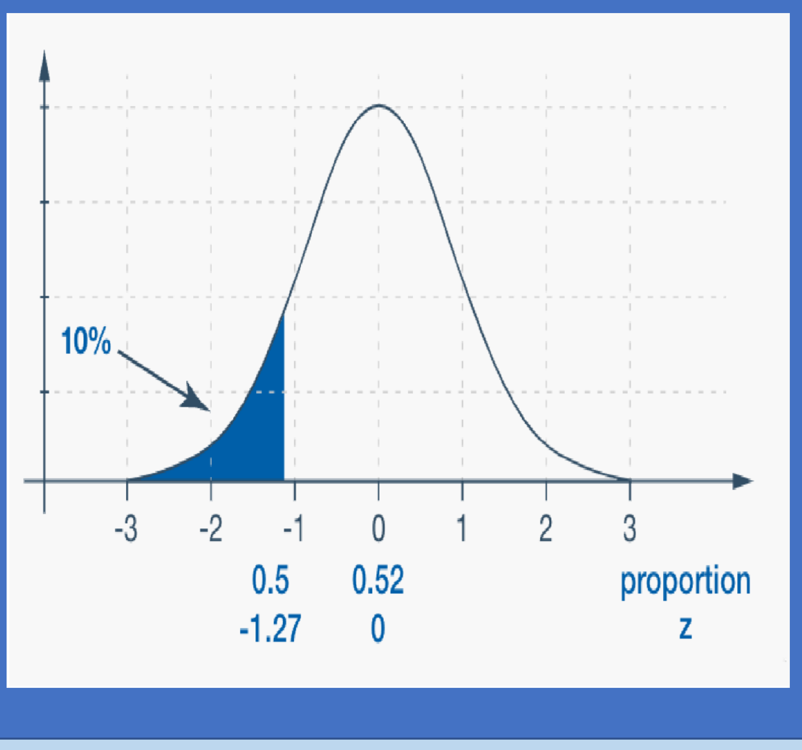
SOLUTION

We know the true population proportion is p = 0.52. So the question is asking about the chances that the sample proportion would come out less than 0.5. The standard deviation of would be:

$$\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.52(0.48)}{1000}} = 0.0158$$

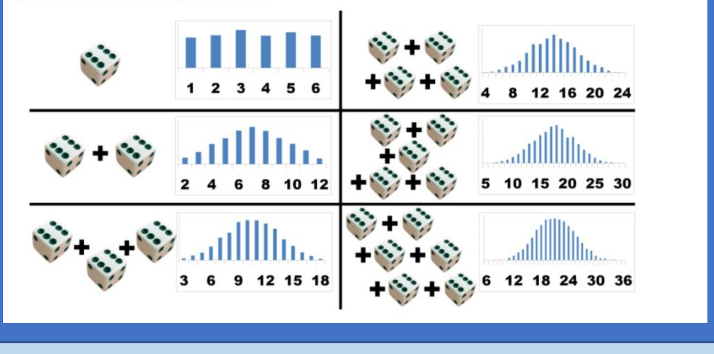
Since the population situation is roughly symmetric (0.52 versus 0.48) the distribution of the sample proportion would follow the normal curve. Thus to compute the probability, we calculate the standard score…

$$z = \frac{(0.5-0.52)}{0.0158} \simeq -1.27$$



EXAMPLE FROM PROBABLITY

If a Dice is rolled, the probability of rolling a one is 1/6, a two is 1/6, a three is also 1/6, etc. The probability of the die landing on any one side is equal to the probability of landing on any of the other five sides. Suppose if there are about 1000 students in a school and each of them is made to the role the same dice then the collection of random variables generated from the above experiment will be sufficiently large and on the histogram, it will tend to become a Normal Distribution.



NORMAL DISTRUBITON TABLE

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952

REFERENCES

[1] Ross Sheldon M. , Introduction to Probablity and Statistics for Engineers and Scientist , Elseveir,California,2014

[2] Akdi Y. , Matematiksel İstatistiğe Giriş,3rd edition,Gazi,Ankara,2011

[3] Bertsekas D.P. and Tsitsiklis J.N. , Introduction to Probablity, 2nd edition,Athene Scientific,U.S.A.,2008

[4] central limit theorem example Archives - Prwatech

[5] Central Limit Theorem Formula | Calculator (Excel Template) (educba.com)

[6] https://www.educba.com/central-limit-theorem-formula/