A. NORMAL DISTRIBUTION

Exercise A1:

The Grear Tire Company has just developed a new steel- belted radial tire that will be sold through a national chain of discounts stores. Because the tire is a new product, Grear's managers believe that the mileage guarantee offered with the tire will be an important factor in the acceptance of the product. Before finalizing the tire mileage guarantee policy, Grear's managers want probability information about the number of miles the tires will last. From actual road tests with the tires, Grear's engineering group has estimated the mean tire mileage is μ= 36.500 miles with a standard deviation of σ = 5.000 miles. In addition the data collected indicate a normal distribution is a reasonable assumption.

1. What percentage of the tires can be expected to last more than 40.000 miles? (24.2%)

Solution

Z=(40.000-36.500)/5.000= 0,7

P(X>40.000)=1-P(X<=40.000) =1-P(Z<=0,7)=1-(0,5+0,2580)=0,242=24,2%

Let us now to assume that Grear is considering a guarantee that will provide a discount on replacement tires if the original tires do not exceed the mileage stated in the guarantee.

1. What should the guarantee mileage be if Grear wants no more than 10% of the tires to be eligible for the discount guarantee? (30.100 miles).

Solution

Z=(X_i-36.500)/5.000. The value of Z for 40% approximately, is 1,28. But is negative because we want the 10% on the left of the mean.

-1,28=(X_i-36.500)/5.000 X_i =30.100 miles

Exercise A2:

The average weekly pay for production workers is 441,84 euros with a standard deviation of 90 euros. Assume that available data indicate that wages are normally distributed.

1. What is the probability that a worker earn between 400 and 500 euros?

Solution

Z_1=(400-441,84)/90=-0,46 Z_2=(500-441,84)/90=0,65

P(400<=X<=500)=P(X<=500)-P(X<=400)=P(Z<=0,65)-P(Z<=-0,46)=

$=\Phi(0,65)-\Phi(-0,46)=(0,5+0,2422)-[1-\Phi(0,46)]=0,7422-1+(0,5+0,1772)=$

$=0,7422-1-0,6772=0,4194=41,94\%$

2. How much does a production worker have to make to be in the top 20% of wage earners?

Solution

$Z=(X_i-441,84)/90$. The value of Z for 30% approximately, is 0,84. It is negative because we want the 20% on the right of the mean.

$0,84=(X_i-441,84)/90 \quad X_i=517,44$ euros.

3. For a randomly selected production worker, what is the probability the worker earns less than 250 euros per week?

Solution

$Z=(250-441,84)/90=-2,13$

$P(X<250)=P(Z<-2,13)=\Phi(-2,13)=1-\Phi(2,13)=1-(0,5+0,4834)=1-0,9834=0,0166=1,66\%$

Exercise A3:

The time needed to complete a final examination in a particular college course is normally distributed with a mean of 80 minutes and a standard deviation of 10 minutes.

1. What is the probability of completing the exam in one hour or less?

Solution

$Z=(60-80)/10=-2$

$P(X<=60)=P(Z<=-2)=\Phi(-2)=1-\Phi(2)=1-(0,5+0,4772)=0,0228=2,28\%$.

2. What is the probability that a student will complete the exam in more than 60 minutes but less than 75 minutes?

Solution

$Z=(60-80)/10=-2 \quad Z=(75-80)/10=-0,5$

$P(60<X<75)=P(X<75)-P(X<60)= P(Z<-0,5)-P(Z<-2)=\Phi(-0,5)-\Phi(-2)=1-\Phi(0,5)-[1-\Phi(2)]=$

$=1-\Phi(0,5)-1+\Phi(2)= \Phi(2)-\Phi(0,5)=0,9772-0,6915=0,2857=28,57\%$.

3. Assume that the class has 60 students and that the examination period is 90 minutes in length. How many students do you expect will be unable to complete the exam in the allotted time?

Solution

Z=(90-80)/10=1

P(X>90)=1-P(X<=90)=1-P(Z<=1)=1-(0,5+0,3413)=0,1587=15,87%

60*0,1587=9,522 about 10 students.

B. INTERVAL ESTIMATION

Exercise B1:

In the testing of a new production method, 18 employees were selected randomly and asked to try the new method. The sample mean production rate for the 18 employees was 80 parts per hour and the sample standard deviation was 10 parts per hour. Provide a 90% and 95% confidence intervals for the population mean production rate for the new method, assuming the population has a normal probability distribution.

Solution

n=18<30, sample mean =80, sample standard deviation=10.

t- student distribution with n-1=18-1=17 d.f

For 90% confidence interval and 17 d.f $t_{0,1/2}=t_{0,05}=1,740$ and the confidence interval is:

$80 \pm 1{,}740 \cdot \frac{10}{\sqrt{18}} = 80 \pm 4{,}1$ 90% (75,9 - 84,1)

For 95% confidence interval and 17 d.f $t_{0,05/2}=t_{0,025}=2,110$ and the confidence interval is

$80 \pm 2{,}110 \cdot \frac{10}{\sqrt{18}} = 80 \pm 4{,}97$ 95% (75,03 - 84,97)

Exercise B2:

In an effort to estimate the mean amount spent per customer for dinner at a major restaurant, data were collected for a sample of 49 customers. Assume a population standard deviation of 5 euros.

1. At 95% confidence, what is the margin of error?

2. If the sample mean is 24,80 euros, what is the 95% confidence interval for the population mean?
3. If de desired range of the confidence interval is 1,5 euros, how large should the sample be at a 95% confidence level?

Solution

n=49, σ = 5

1. At 95% confidence level the value of z is 1.96.

$$E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1{,}96 \cdot \frac{5}{\sqrt{49}} = 1{,}4$$

2. Sample mean=24,8

Confidence interval: $\bar{x} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 24{,}8 \pm 1{,}96 \cdot \frac{5}{\sqrt{49}} = 24{,}9 \pm 1{,}4$

95% (23,5 - 26,3)

3. The sample size should be: $n = \frac{(2 \cdot Z_{\alpha/2})^2 \cdot \sigma^2}{R^2} = \frac{(2 \cdot 1{,}96)^2 \cdot 5^2}{1{,}5^2} = 170{,}74 \approx 171$

Exercise B3:

Audience profile data collected at the ESPN SportsZone web site showed that 104 of the users were women. Assume that this percentage was based on a sample of 400 users.

1. Using 95% confidence, what is the margin of error associated with the estimated proportion of users who are women?
2. What is the 95% confidence interval for the population proportion of ESPN SportZone web site users who are women?
3. How large a sample should be taken if the desired margin of error is 3%?.

Solution

$\bar{p} = \frac{104}{400} = 0{,}26$ $1 - \bar{p} = 1 - 0{,}26 = 0{,}74$ $Z_{0,05/2} = Z_{0,025} = 1{,}96$

1. $E = Z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = 1{,}96 \cdot \sqrt{\frac{0{,}26 \cdot 0{,}74}{400}} = 0{,}043$

2. $\bar{p} \pm Z_{\alpha/2} \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = 0{,}26 \pm 0{,}043$ 95% (0,217 - 0,303) 21,7% - 30,3%

3. The proportion of the population is unknown and thus, if we assume that the sample proportion is about the same, we can calculate the sample size approximately.

$$n = \frac{(z_{a/2})^2\, p(1-p)}{E^2} = \frac{1{,}96^2 \cdot 0{,}26 \cdot 0{,}74}{0{,}03^2} = 821$$

Exercise B4:

An Associated Press poll of 1018 adults found 255 adults planned to spend less money on gifts during the 2018 holiday season compared to the previous year.

1. Using 95% confidence level, what is the margin of error associated with this estimate?
2. What is the 95% confidence interval for the population proportion?
3. How large a sample should be taken if the desired range is 4%?.

Solution

$\overline{p} = \frac{255}{1018} = 0{,}25$ $\quad 1 - \overline{p} = 1 - 0{,}25 = 0{,}75$ $\quad Z_{0,05/2} = Z_{0,025} = 1{,}96$

1. $E = Z_{a/2}\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 1{,}96 \cdot \sqrt{\frac{0{,}25 \cdot 0{,}75}{1018}} = 0{,}0266$

2. $\overline{p} \pm Z_{a/2}\sqrt{\frac{\overline{p}(1-\overline{p})}{n}} = 0{,}25 \pm 0{,}0266$ 95% (0,2234 - 0,2766) 22,34% -27,66%

3. The proportion of the population is unknown and thus, if we assume that the sample proportion is about the same, we can calculate the sample size approximately.

$$n = \frac{(z_{a/2})^2\, p(1-p)}{E^2} = \frac{1{,}96^2 \cdot 0{,}25 \cdot 0{,}75}{0{,}04^2} = 450$$

Exercise B5:

The U.S Department of Transportation provides the number of miles that residents of the 75 largest metropolitan areas travel per day in a car. Suppose that for a simple random sample of 50 Buffalo residents the mean is 22,5 miles a day and the standard deviation is 8,4 miles a day, and from the independent simple random sample of 100 Boston residents the mean is 18,6 miles a day and the standard deviation is 7,4 miles a day.

What is the 95% confidence interval for the difference between the two population means?

Solution

Buffalo n_1=50 mean=22,5 miles st.d =8,4 miles

Boston n_2=100 mean=18,6 miles st.d =7,4 miles

Large samples with σ_1 and σ_2 unknown. Significance level =0,05 and $Z_{0,05/2}$=$Z_{0,025}$ =1,96

$$(\bar{x}_1 - \bar{x}_2) \pm Z_{a/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (22,5 - 18,6) \pm 1,96 \cdot \sqrt{\frac{8,4^2}{50} + \frac{7,4^2}{100}}$$

$$= 3,9 \pm 1,96 * 1,399$$

95% (1,157 - 6,642)

Exercise B6:

An urban planning group is interested in estimating the difference between the mean household incomes for two neighborhoods in a large metropolitan area. Independent random samples of households in the neighborhoods provided the following results.

Neighborhood 1	**Neighborhood 2**
n_1 =8	n_2 =12
Mean 1 =15.700$	Mean 2 =14.500$
St.d 1=700$	St.d 2=850$

1. Develop a 95% confidence interval for the difference between the mean incomes in the two neighborhoods.
2. What assumptions were made to compute the interval estimates in part 1?.

Solution

Small samples with Standard deviations σ_1 and σ_2 unknown. Thus, we should use the pooled estimator of σ^2.

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(8-1)700^2 + (12-1)850^2}{8+12-2} = 632083,33$$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})} = \sqrt{632083,33(\frac{1}{8} + \frac{1}{12})} = 362,88$$

The value of $t_{a/2} = t_{0,025}$ with 8+12-2=18 d.f is 2,101.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{a/2} \cdot \sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})} = (15700 - 14500) \pm 2,101 \cdot 362,88 = 1200 \pm 762,41$$

95% (437,59 - 1962,41)

Exercise B7:

A sample of 6 cars was selected at random and type A gasoline was placed in each of them and the distance traveled was measured. Subsequently, the same cars were loaded with the same amount of Type B gasoline and the distance they traveled was again measured. Assuming that the tests were performed under the same standard conditions, the results are given in the table (in Km).

Car	**1**	**2**	**3**	**4**	**5**	**6**
Type A	125	64	94	38	90	106
Type B	133	65	103	37	102	115

Provide a 95% confidence interval for the difference in gasoline consumption.

Solution

Two Dependent Samples. The same cars with two different gasoline types.

Car	**1**	**2**	**3**	**4**	**5**	**6**	
Type A	125	64	94	38	90	106	
Type B	133	65	103	37	102	115	
d_i	8	1	9	-1	12	9	**38**
d_i^2	64	1	81	1	144	81	**372**

$$\bar{d} = \frac{\sum d_i}{n} = \frac{38}{6} = 6{,}33 \qquad s_d = \sqrt{\frac{\sum d_i^2 - \left[\frac{(\sum d_i)^2}{n}\right]}{n-1}} = \sqrt{\frac{372 - \frac{38^2}{6}}{6-1}} = 5{,}1$$

$$\bar{d} \pm t_{a/2} \cdot \frac{s_d}{\sqrt{n}} = 6{,}33 \pm 2{,}57 * \frac{5{,}1}{\sqrt{6}} = 6{,}33 \pm 5{,}35$$

Lower limit: 0,98 Upper Limit: 11,68

Exercise B8:

Independent simple random samples of tax returns from two offices provide the following information.

Office 1	**Office 2**
n_1=250	n_2= 300

Number of returns with errors =35	Number of returns with errors =27

Provide a 95% confidence interval for the difference between the two population error rates.

Solution

$$\bar{p}_1 = \frac{x}{n_1} = \frac{35}{250} = 0{,}14 \qquad \bar{p}_2 = \frac{x}{n_2} = \frac{27}{300} = 0{,}09$$

$$s_{\bar{p}_1-\bar{p}_2} = \sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}} = \sqrt{\frac{0{,}14*0{,}86}{250} + \frac{0{,}09*0{,}91}{300}} = 0{,}0275$$

$$(\bar{p}_1 - \bar{p}_2) \pm Z_{\alpha/2} \cdot \sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}} = (0{,}14 - 0{,}09) \pm 1{,}96 * 0{,}0275 = 0{,}05 \pm 0{,}0539$$

C. HYPOTHESES TESTING

Exercise C1:

In a study of driving practices, it was found that 48% of drivers did not stop at stop sign intersections on country roads. Assumed that a follow –up study two months later found the 360 of 800 drivers did not stop at stop sign intersections on country roads. The proportion of drivers who did not stop has changed? What is your conclusion? (a=5%). At what level of significance would you change your decision?

Solution

p=0,48 1-p=1-0,48=0,52 $\bar{p}$=360/800=0,45 1- $\bar{p}$=1-0,45=0,55

H_0: p=0,48

H_1: p $\neq$ 0,48

$$\sigma_{\bar{p}} = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{0{,}48 \cdot 0{,}52}{800}} = 0{,}0176 \qquad Z = \frac{(\bar{p} - p_0)}{\sigma_{\bar{p}}} = \frac{0{,}45 - 0{,}48}{0{,}0176} = -1{,}70$$

$Z_{0,05/2}$=$Z_{0,025}$=1,96 $|-1{,}70| = 1{,}70 < 1{,}96$

The proportion of drivers who did not stop did not change.

Our opinion can be changed if $Z_{a/2}$ is less than 1,70. For example $Z_{a/2}$=1,69 and from the table of Normal Distribution the probability that corresponds is 0,4545.

0,5-0,4545=0,0455*2=0,091=9,1%

Exercise C2:

Past years the mean length of a work week for the population of workers was 39,2 hours. This year a random sample of 112 workers showed a mean of 38,5 hours and a standard deviation of 4,5 hours. Do you believe that the mean length of a work week is less than 39,2 hours? (a= 0,05).

Solution

μ_0=39,2 n=112 sample mean= 38,5 s=4,5 a=0,05

H_0: μ=39,2

H_1: μ <39,2

$$Z = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(38{,}5 - 39{,}2)\sqrt{112}}{4{,}5} = -1{,}65$$

$\mathbf{Z}_{0,05}$= 1,645 we reject H_0 if $Z<-Z_a$ Z= -1,65<-1,645= $\mathbf{Z}_{0,05}$ and thus we can conclude that the mean length of a work week is less than 39,2 hours.

Exercise C3:

The average U.S household spends $90 per day. A sample of 25 households, in Corning, New York, showed a sample mean daily expenditures of $84,5 with a sample standard deviation of $14,5. What you can say about the household spends in Corning? (a=5%).

Solution

μ_0=90 n=25 sample mean= 84,5 s=14,5 a=0,05

H_0: μ=90

H_1: μ≠90

$$t = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(84{,}5 - 90)\sqrt{25}}{14{,}5} = -1{,}89$$

$t_{0,05/2}$=2,064 with 25-1=24 d.f Rejection Rule: $|t| > t_{a/2}$

In this case $|-1{,}89| < 2{,}064$ and thus we accept the null hypothesis.

Exercise C4:

A production line operates with a mean filling weight of 455 grams per container. Overfilling or underfilling is a serious problem, and the production line should be shut down if either occurs. From past data, σ is assumed to be 24 grams. A quality control inspector samples 36 items every two hours and at the time makes decision of whether to shut the line down for adjustment.

1. With a 0,05 level of significance, what is the rejection rule for the hypothesis testing procedure?
2. If a sample mean of 462 grams were found, what action would you recommended?
3. If a sample mean of 449 grams were found, what action would you recommended?

Solution

μ_0=455   n=36   σ=24

1. H_0: μ=455

   H_1: μ≠455

With a 0,05 level of significance, the rejection rule for the hypothesis testing procedure is:

$|Z| > Z_{a/2}$

2. H_0: μ=455

   H_1: μ≠455

$$Z = \frac{(\bar{x} - \mu_0)\sqrt{n}}{\sigma} = \frac{(462 - 455)\sqrt{36}}{24} = 1{,}75$$

The rejection rule is: $|Z| > Z_{a/2}$

$Z_{0,05/2}=Z_{0,025}$= 1,96 $|1{,}75| < 1{,}96$

We would recommend to continue the production line.

3. H_0: μ=455

   H_1: μ≠455

$$Z = \frac{(\bar{x} - \mu_0)\sqrt{n}}{\sigma} = \frac{(449 - 455)\sqrt{36}}{24} = -1{,}5$$

The rejection rule is: $|Z| > Z_{a/2}$

$$|-1{,}5| < 1{,}96$$

We would recommend to continue the production line.

To verify the above results we can construct the corresponding confidence intervals for sample means of 462 and 449.

For sample mean 462 the confidence interval is: $\bar{x} \pm Z_{a/2} \cdot \frac{\sigma}{\sqrt{n}} = 462 \pm 1{,}96 \frac{24}{\sqrt{36}} = 462 \pm 7{,}84$

(454,16 – 469,84) In the confidence interval belongs the desired value of 455 and thus the null hypothesis is not rejected.

For sample mean 449 the confidence interval is: $\bar{x} \pm Z_{a/2} \cdot \frac{\sigma}{\sqrt{n}} = 449 \pm 1{,}96 \frac{24}{\sqrt{36}} = 449 \pm 7{,}84$

(441,16 – 456,84). For the same reason the null hypothesis is not rejected.

Exercise C5:

Suppose a new production method will be implemented if a hypothesis test supports the conclusion that the new method reduces the mean operating cost per hour. A quality control inspector took 24 samples, one per hour, and the mean cost for the new production method was 218 euros per hour with a standard deviation of 4 euros. State the appropriate null and alternative hypotheses if the mean cost for the current production method is 220 euros per hour.

Solution

μ_0=220   n=24   sample mean= 218   s=4 a=0,05

H_0: μ=220

H_1: μ<220

$$t = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(218 - 220)\sqrt{24}}{4} = -2,45$$

$t_{0,05}$=1,714 with 24-1=23 d.f   Rejection Rule: $t < -t_a$

In this case   t=-2,45<-1,714= $t_{0,05}$ and thus we reject the null hypothesis. That means that the new method reduces the mean operating cost per hour.