Recitation 3

carlo.cavicchia@uniroma1.it

Exercise 1

The Mental Development Index (MDI) of the Bayley Scales of Infant Development is a standardized measure used in observing infants over time. It is approximately normal with a mean of 100 and a standard deviation of 16.

1. What proportion of children has an MDI of
   (i) at least 120?
   (ii) at least 80?
   Answer
   (i) $P(X > 120) = P(\frac{X-100}{16} > \frac{120-100}{16}) = P(Z > 1.25) = P(Z < -1.25) = 0.1056$

   (ii) $P(X > 80) = P(\frac{X-100}{16} > \frac{80-100}{16}) = P(Z > -1.25) = 1 - P(Z < -1.25) = 1 - 0.1056 = 0.8944$

2. Compute the probability that in a sample of: (ii) 25 babies, the proportion of children with an MDI is greater than 90;

   Answer
   $P(\bar{X} > 90) = P(\frac{\bar{X}-100}{16/\sqrt{25}} > \frac{90-100}{16/\sqrt{25}}) = P(Z > -3.13) = 1-P(Z < -3.13) = 0.9991$

Exercise 2

The average monthly rental rate for one-bedroom apartments in Rome is 650 euros. Suppose rental rates across all one-bedroom apartments in Rome follow approximately a normal distribution, with a standard deviation of 100 euros.

1. Find the approximate proportion of one-bedroom apartments for which the rental rate is between 500 euros and 900 euros a month.
   Answer

$P(500 < X < 900) = P(\frac{500-650}{100} < \frac{X-650}{100} < \frac{900-650}{100}) = P(-1.5 < Z < 2.5) =$
$\Phi(2.5) - \Phi(-1.5) = 0.9938 - 0.0668 = 0.927$

2. Compute the probability that in a sample of 5 one-bedroom apartments, the proportion of one-bedroom apartments for which the rental rate is less than 500 euros a month;
   Answer
   $P(\bar{X} < 500) = P(\frac{\bar{X}-650}{100/\sqrt{5}} < \frac{500-650}{100/\sqrt{5}}) = P(Z < -3.35) = 0.0004$

Exercise 3

It has been asked "What do you think is the ideal number of children for a family to have?". The 50 females who responded had a median of 2, mean of 3.22, and standard deviation of 1.99.

1. What is the point estimate of the population mean?
   Answer
   The sample mean.

2. Is the estimator used in (a) unbiased?

   Answer
   It is unbiased. Its standard error tends to 0 when the sample size goes to inifinity.

3. Find the standard error of the sample mean.
   Answer
   $1.99/\sqrt{50} = 0.2814$

4. Test the null hypothesis $H_0 : \mu = 3$ vs the alternative $H_1 : \mu > 3$ at the significance level $\alpha = 0.05$.
   Answer

   Step 1: Assumptions
   The variable is quantitative, the sample size is large enough.

   Step 2: Hypotheses
   $H_0 : \mu = 3$ vs the alternative $H_1 : \mu > 3$
   Step 3: Test statistic
   $z_{oss} = \sqrt{n} \times (\bar{x} - \mu_0)/\sigma = \sqrt{50} \times (3.22 - 3)/1.99 = 0.7817$
   Step 4: P-value
   p-value $= P(Z > z_{oss}) = P(Z > 0.7817) = 0.2172$

   Step 5: Conclusion
   p-value $> \alpha$. Thus we do not reject H_0

Exercise 4

Suppose the mean GPA of all students graduating from Penn State University in 2005 was 3.05. The registrar plans to look at records of 100 students graduating in 2015 to see if mean GPA has changed. The sample mean is 2.15, while standard deviation is 1.5.

1. State the null and alternative hypotheses for this investigation.
   Answer
   Test the null hypothesis $H_0 : \mu = 3.05$ vs the alternative $H_1 : \mu \neq 3.05$

2. Test the null hypothesis H_0 vs the alternative H_1 at the significance level $\alpha = 0.05$.

   Answer

   Step 1: Assumptions
   The variable is quantitative, the sample size is large enough.

   Step 2: Hypotheses
   $H_0 : \mu = 3.05$ vs the alternative $H_1 : \mu \neq 3.05$

   Step 3: Test statistic
   $z_{oss} = \sqrt{n} \times (\bar{x} - \mu_0)/\sigma = \sqrt{100} \times (2.15 - 3.05)/1.5 = -6$
   Step 4: P-value
   p-value $= P(Z > |z_{oss}|) = 2P(Z > 6) \approx 0$

   Step 5: Conclusion
   p-value $< \alpha$. Thus we do reject H_0

Exercise 5

Here it is the distribution of family year income (ten thousands of dollars) in a small village: $4, 5, 6, 5, 3, 10, 5, 8, 4, 7$.

1. What type of statistical variable is it?

   Answer
   Quantitative continuous

2. Find mean and mode.
   Answer
   Mean=5.7; Median=5.

3. Is the distribution of these data symmetric?
   Answer
   The mean is different from the median. The mean is equal to 5.7, while the median is

equal to 5. It follows that the distribution is skewed to the right, since the median is smaller than the mean (in other words the distribution has a long right tail).

4. Compute the variance.
   Answer

   The variance is 4.01.

   $\sigma^2 = (2\times(4-5.7)^2+3\times(5-5.7)^2+(6-5.7)^2+(3-5.7)^2+(7-5.7)^2+(8-5.7)^2+(10-5.7)^2)/10 = 4.01$

Exercise 6

The following table contains income (X) and spending (Y) amounts for 5 families:

x	y
6	5
7	6
9	7
5	4
8	5

1. Are the variable correlated? Compute the correlation coefficient and interpret it.

   Answer
   $\bar{x} = 7$, $\bar{y} = 5.4$

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x}) \times (y - \bar{y})$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
6	5	-1	-0.4	0.4	1	0.16
7	6	0	0.6	0	0	0.36
9	7	2	1.6	3.2	4	2.56
5	4	-2	-1.4	2.8	4	1.96
8	5	1	-0.4	-0.4	1	0.16
				6	**10**	**5.2**

   It follows that $r = 6/\sqrt{(10 \times 5.2)} = 0.83$.
   The correlation is positive and quite strong. It follows that income and spending are lineraly dependent.