

Examples

Marius Ionescu

11/29/2011

Example

Example

100,000 tax forms are reported to have an average income of \$12,000 with an SD of \$6000. Additional study of 900 forms is proposed. What is the chance that income on these 900 forms will average between \$11,800 and \$12,200? $EV_{AVG} = 12,000$; $SE_{AVG} = \frac{\sqrt{900 \cdot 6000}}{900} = 200$; 68%

Example

Example

Suppose that a sample of 12714 people were asked about their voting preferences and if they voted or not. The results are as follows:

	Voted	Didn't vote	Totals
Dem	3113	1512	4625
Rep	2148	1004	3152
Ind	2950	1423	4373
Other	111	453	564
Totals	8322	4392	12714

Is there a difference in the voting preferences between the people who voted and those that did not?

Solution

	%	Voted	Didn't vote
Dem	36%	$36\% \cdot 8322 = 2995.9$	1581.1
Rep	25%	2080.5	1098
Ind	34%	2829.5	1493.3
Other	5%	416.1	219.6
Totals	100%		

$$\chi^2 = 277.62$$

degrees of freedom = $(4-1) \cdot (2-1) = 3$; $P = 0\%$

Example

Example

A fair die is rolled four times. What is the chance of getting a 1 or a 2 at least once?

Example

Example

Testing the effect of vitamin A on rat learning: 800 rats are paired at random. One of each pairs gets vitamin A supplements, then both run a maze. The quantity

untreated rat's time - treated rat's time

has an average of 1 sec, with an SD of 5 sec. Did vitamin A help the rats to learn to run the maze? Or was it just chance variation?

Null hypothesis: difference is 0; observed is 1;

$$SE_{AVG} = \frac{SD_{sample}}{\sqrt{n}} = \frac{5}{\sqrt{400}} = 0.25; Z = \frac{1-0}{0.25} = 4; P \simeq 0;$$

Example

Example

740 Colgate students take 32 courses in 4 years. Suppose grades are given only with letter values (no +/-) and the numbers 0,1,2,3,4 are assigned to F,D,C,B,A. Let's test the idea that professors give out grades randomly.

- If grades are assigned randomly, how many students do we expect to have GPA of 3.0 or higher? $AVG_{box} = 2$; $SD_{box} = 1.41$; $EV_{AVG} = 2$; $SE_{AVG} = \frac{\sqrt{32} \cdot 1.41}{32} = 0.25$; $Z=4$; $P(X \geq 3) = \frac{100 - A(4)}{2} \simeq 0.003\%$; out of 740 students: $\simeq 0.02$ students; basically no honor students at Colgate

Example

Example

A box containing two 1's and one 2 has two items drawn from it without replacement. What is the chance that the second item drawn is a 1?

Example

Example (Adapted from the Syracuse Post-Standard, Dec. 9, 1999)

Swedish scientists examined all births in Sweden between 1987 and 1995, and found that 1 million children were born in that time, 51% of them girls. But of the 5900 women who were admitted to the hospital during their first trimester with vomiting and extreme nausea, 56% had girls. If a woman has severe morning sickness in her first trimester, is she more likely to have a girl?

Null hypothesis: no, the draw is from the same box with average 51%; the difference is due to chance. Box 51% of 1s and 49% of 0s. $EV_{\%} = 51\%$;

$$SD_{box} = (1 - 0)\sqrt{\frac{51}{100} \times \frac{49}{100}} = 0.499;$$

$$SE_{\%} = \frac{4.99}{\sqrt{5900}} = 0.0065 * 100\% = 0.65\%; \quad Z = \frac{56-51}{0.65} \simeq 7.7; \quad P = 0$$

Example

Example

A class consists of three freshmen, six sophomores, and 11 juniors. Three members are selected at random (without replacement). Find the chance that

- The second one chosen is a sophomore.
- The second is a sophomore and the third is a junior.
- The second and third are both sophomores.

Example

Example

A standard math test is given to 25 randomly chosen Colgate students and 25 Hamilton College students. The Colgate students averaged 155 of a possible 200, with an SD of 15; the Hamilton students averaged 150, with an SD of 20. Are Colgate students better at math?

null hypothesis: the two schools are the same; observed

difference=155-150=5; $SD_{AVG1} = \frac{15}{\sqrt{25}} = 3$; $SD_{AVG2} = \frac{20}{\sqrt{25}} = 4$;

$SE_{diff;AVG} = \sqrt{3^2 + 4^2} = 5$; $Z = 1$; $P \simeq 16\%$; can not reject the null hypothesis;

Example

Example

Does giving the drug AZT to HIV-positive pregnant women save their babies from HIV? (Newsweek, March 7, 1994) Out of 163 babies born to mothers treated with AZT, only 13 were HIV-positive, while out of 161 born to mothers treated with a placebo, 40 were HIV-positive. null hypothesis: the two groups are the same; % difference = 0; observed difference = $\frac{13}{163} - \frac{40}{161} = -0.17 * 100\% = -17\%$;

$$SD_{box1} = \sqrt{\frac{13}{163} \frac{150}{163}} = 0.27; SD_{box2} = \sqrt{\frac{40}{161} \frac{121}{161}} = 0.43;$$

$$SE_{1\%} = \frac{0.27}{\sqrt{163}} 100\% = 2.11\%; SE_{2\%} = \frac{0.43}{\sqrt{161}} 100\% = 3.39\%;$$

$$SE_{diff\%} = \sqrt{2.11^2 + 3.39^2} = 3.99; Z = \frac{-17}{3.99} = -4.26; P \simeq 0\%; \text{reject the null hypothesis}$$

Example

Example (Poker hands)

Find the probability of the following poker hands:

- two pair $\frac{(13\ 2)(4\ 2)^2(11\ 1)(4\ 1)}{(52\ 5)}$
- full house $\frac{(13\ 1)(4\ 3)(12\ 1)(4\ 2)}{(52\ 5)}$
- four of a kind $\frac{(13\ 1)(4\ 3)(12\ 1)(4\ 2)}{(52\ 5)}$

Example

Example

Improvements are made to a drug for cold symptoms, and it is desired to test whether it lasts for more than 2 hours. Fifteen subjects colds are given the drug, and it is found that its effects wear off after 2.2 hours with an SD of .3 hours. Does the drug really last for more than 2 hours (on the

average)? t-test: $AVG_{box} = 2$; $SD_{box} = \sqrt{\frac{15}{14}} \cdot .3 = 0.31$; $EV_{AVG} = 2$;
 $SE_{AVG} = \frac{0.31}{\sqrt{15}} = 0.08$; $t = \frac{2.2-2}{0.08} = 2.5$; degree of freedom=14; $p = 1\%$

Example

Example (Is there a prenatal basis for homosexuality?)

(McFadden and Pasanen, Proc. Natl. Acad. Sci. USA 95 (1998)). If a quiet click sound is made outside a person's ear, the inner ear responds with "otoacoustic emissions" (OAEs), very weak vibrations that can be measured by a microphone in the ear canal. The inner ears of women usually generate stronger (higher amplitude) OAEs than those of men in response to clicks of the same strength; this is attributed to the androgens which affect the male fetus in the womb. In this study, the right ears of 57 heterosexual women produced OAEs of average amplitude 18.2 SPL, with an SE of 0.8 dB, in response to a click of 75dB; the corresponding figures for 37 homosexual women was 16.0 dB and 0.7 dB. Is there a significant difference in the OAEs of the homosexual vs. heterosexual women? z-test: $EV_{diff} = 0$; $Actual_{diff} = 18.2 - 16 = 2.2$; $SE_{diff} = \sqrt{0.8^2 + 0.7^2} = 1.06$; $Z = \frac{2.2}{1.06} = 2.07$; $p = 2\%$

Example

Example

The African kumquat has short, intermediate and long maturation times. One genetic model says that maturation time is controlled by a single gene, of which there are two forms, short (*s*) and long (*l*), neither dominant, so that the intermediate maturation time appears in the genotype *s/l* (or *l/s*). Suppose some short-maturing and some intermediate-maturing kumquats are crossed, and of the 20 progeny selected at random, 7 are short-maturing and 13 are intermediate. Does that fit the model?

		<i>s</i>	<i>l</i>	Observed	
Model	<i>s</i>	1/4	1/4	7	10
	<i>l</i>	1/4	1/4	13	10

$\chi^2 = \frac{(7-10)^2}{10} + 1.8$; degree of fr=1; p=20%