Chapter One

Ryan B. Honea

Exercise One

Question

As in Example 1.1, suppose that one machine produces rods with diameters normally distributed with mean 27 mm and standard deviation 1.53 mm, so that 2.5 percent of the rods have diameter 30 mm or more. A second machine is known to produce rods with diameters normally distributed with mean 24 mm and 2.5 percent of rods it produces have diameter 30 mm or more. What is the standard deviation of rods produced by the second machine?

Solution

We will solve this by utilizing the formula for z-score. The upper 2.5% of data is located 1.96 standard deviations to the right of the mean or at $z = 1.96$. The formula for z-score is

$$z_x = \frac{x - \mu}{\sigma}$$

where x is the value being considered. In this case, we have

$$1.96 = \frac{30 - 24}{\sigma} \to \sigma = \frac{30 - 24}{1.96} \to \sigma \approx 3.06$$

So, the standard deviation is approximately 3.06.

Exercise Two

Question

In a group of 145 patients admitted to a hospital with a stroke, weekly alcohol consumption in standard units had a mean of 17 and a standard deviation of 22. Explain why their alcohol consumption does not follow a normal distribution. Is this finding surprising?

Solution

In a normal distribution, approximately 68% of the data lies between one standard deviation from the mean. Therefore, if this group were normally distributed, 68% of the patents would have a weekly alcohol consumption of -5 units to 39 units which does not make sense as -5 units is an impossible value of weekly consumption. This is likely a right-skewed distribution which isn't altogether surprising. Though this based on experience, I would argue the number of binge drinkers is low, but the volume of alcohol a binge drinker might consume is enough to skew the standard deviation.

Exercise Three

Question

Following a television campaign about the risks of smoking tobacco, the cigarette consumption of a group of 50 smokers decreases by a mean of 5 cigarettes per day with a standard deviation of 8. Explain why the reasoning of Exercise 2 cannot be used to show that this distribution is not normal.

Solution

In this case, we aren't sure of the current amount of cigarettes our sample smokes. The argument of Exercise 2 relies on knowing a minimum where in this case a smokers habit could change from between -13 to +3 cigarettes a day. Even out to three standard deviations, it's completely reasonable to believe that a smoker might change -29 and +19 cigarettes a day in the most extreme of cases. There is no evidence to suggest that it isn't normal.

Exercise Four

Question

In section 1.3, we pointed out that 5 tosses of a coin would never provide evidence against the hypothesis that a coin was fair (equally likely to fall heads or tails) at a conventional 5 percent significance level. What is the least number of tosses needed to provide such evidence using a two-tail test, and what is then the exact P-value.

Solution

Following the logic of section 1.3 which takes the most extreme case (which would give the least number of tosses needed), we are interested in finding the probability that all the tosses are either all heads or all tails. So, we want to find x where

$$P = .05 >= 2 * (0.5)^x$$

Note that this is essentially a two-tailed test.

$$.025 >= (0.5)^x \rightarrow x = \log_{0.5}(0.025) \approx 5.32$$

To find the minimum number of tosses, we take the ceiling of that result which is six tosses. The exact P-value it would give is $P = 2 * (0.5)^6 = .03125$.

Exercise Five

Question

A biased coin is such that Pr(heads) $= 2/3$. If this coin is tossed, the least number of times calculated in Exercise 1.4, what is the probability of an error of the second kind associated with the 5 percent significance level? What is the power of the test? Does the discrete nature of possible P-values cause any problems in calculating the power?

Solution

Consider this a two-tailed binomial test where p determines the probability of flipping a heads. Then the tests are $H_0 : p = .5$ and $p : \mu \neq .5$. The sample probability interval in which we wouldn't reject the null hypothesis for this test would therefore come from the binomial confidence interval equation

$$p \pm z_{\alpha/2}\sqrt{\frac{1}{n}p(1-p)} \rightarrow .5 \pm 1.65\sqrt{\frac{1}{6}.5 * .5} \rightarrow (.163, .834)$$

In terms of coins flipped, that interval then means that we should observe at least 1 head and at most 5 heads in order to not reject the null hypothesis. If the probability that heads is flipped is actually 2/3, then we can use the binomial cumulative distribution to find the probability that we will fail to reject the null hypothesis which would result in type two error. I use R to determine this value.

```
pbinom(5, size = 6, prob = (2/3)) - pbinom(0, size = 6, prob = (2/3))
```

```
## [1] 0.9108368
```

The Type II error therefore is approximately .911 while the power of the test is .089 which is not a strong value.

The discrete nature of P here causes slight problems because the typical method for calculating power would use the critical probabilities which are continuous while the number of heads is discrete. If you could observe fractions of heads, then the actual interval of heads is $(.978, 5.004)$. That being said, binomials don't work that way and so the current power of the test is the most accurate result possible and not problematic.

Exercise Six

Question

If a random variable X_i is distributed $\mathrm{N}(\mu, \sigma^2)$ and all X_i are independent, it is well know that the variable

$$Y = \sum_{i=1}^{n} X_i$$

is distributed $N(n\mu, n\sigma^2)$. Use this result to answer the following:

The times in minutes a farmer takes to place any fence post are each independently distributed $\mathrm{N}(10, 2)$. He starts placing posts at 9 a.m one morning, and immediately after one post is placed he proceeds to place another, continuing until he has placed 9 posts. What is the probability that he has placed all 9 posts by (i) 10.25 a.m, (ii) 10.30 am, and (iii) 10.40 a.m?

Solution

It will be easier to answer this in terms of minutes past instead of by time. The problems will instead be considered as (i) at most 85 minutes, (ii) at most 90 minutes, and (iii) at most 100 minutes. Based on the information given, the minutes he completes the task in is distributed $\mathrm{N}(90, 18)$. For each subproblem, I will use the normal cdf in R.

```
pnorm(85, 90, sqrt(18))
```

```
## [1] 0.1192964
```

```
pnorm(90, 90, sqrt(18))
```

```
## [1] 0.5
```

```
pnorm(100, 90, sqrt(18))
```

```
## [1] 0.9907889
```

`pnorm(x,mu,std)` calculates $P(X < x)$ where mu is your mean and std is your standard deviation. The answers therefore are: (i) .119, (ii) .500, and (iii) .991

Exercise Seven

Question

The following two sample data sets both have sample mean 6.

Set I	13.9	2.7	0.8	11.3	1.3
Set II	2.7	8.3	5.2	7.1	6.7

Set I	Set II
13.9	2.7
2.7	8.3
0.8	5.2
11.3	7.1
1.3	6.7

If μ is the population mean, perform for each set t-tests of (i) $H_0 : \mu = 8$ against $H_1 : \mu \neq 8$ and (ii) $H_0 : \mu = 10$ against $H_1 : \mu \neq 10$. Do you consider the conclusions of the tests reasonable? Have you any reservations about using a t-test for either of these data sets?

Solution

I'll address my one reservation before performing the tests. One of the assumptions that must be made when performing a t-test is a reasonably large sample size. A sample size of 5 would not typically be considered a large sample size. The results of the test are as follows:

```r
SetI <- c(13.9, 2.7, 0.8, 11.3, 1.3)
SetII <- c(2.7, 8.3, 5.2, 7.1, 6.7)
mu_i <- 8
mu_ii <- 10
t.test(SetI, mu = mu_i)$p.value
```

```
## [1] 0.5063734
```

```r
t.test(SetII, mu = mu_i)$p.value
```

```
## [1] 0.1062178
```

```r
t.test(SetI, mu = mu_ii)$p.value
```

```
## [1] 0.2185681
```

```r
t.test(SetII, mu = mu_ii)$p.value
```

```
## [1] 0.0141824
```

The only case where the P-value is lower than 5% is whether SetII is different from $\mu = 10$ so we reject H_0 in that case. All of the others, we fail to reject the null hypothesis.

Exercise Eight

Question

Use an available standard statistical software package, or one of the many published tables of binomial probabilities to determine, for sample of 12 from binomial distributions with $p = 0.5$ and with $p = 0.75$, the probabilities of observing each possible number of outcomes for each of these values of p. In a two-tail test of the hypotheses, $H_0 : p = 0.5$ against $H_1 : p = 0.75$ what is the largest attainable P-value less than 0.05? What is the critical region for a test based on this P-value? What is the power of the test?

Solution

Below is the table representing the number of successes or each possible number of outcomes.

Outcomes	50%	75%
0	0.0002441	5.96e-08
1	0.00293	2.146e-06
2	0.01611	3.541e-05
3	0.05371	0.0003541
4	0.1208	0.00239
5	0.1934	0.01147
6	0.2256	0.04015
7	0.1934	0.1032
8	0.1208	0.1936
9	0.05371	0.2581
10	0.01611	0.2323
11	0.00293	0.1267
12	0.0002441	0.03168

Based on these p-values, we can see that the P-values less than 0.5 are 0 to 2 and 10 to 12. The sum of these is approximately 0.0386, i.e. our max P-value. The power of our test would be obtaining a result within this critical region, or 1 minus the probability of obtaining a result outside of it.

```
1 - (pbinom(9, size = 12, prob = 0.75) -
     pbinom(2, size = 12, prob = 0.75))
```

```
## [1] 0.3907126
```

So the power of this test is approximately 0.391