Classical and Bayesian Statistics

Problems 4

Problem 4.1

($**$) A wine merchant claims that the wine bottles he fills contain 70 centiliters. However, a sceptical consumer suspects that the wine dealer is bottling too little wine and wants to verify this claim. He therefore buys 12 bottles of wine and measures their contents. He finds:

$$71, 69, 67, 68, 73, 72, 71, 71, 68, 72, 69, 72 \quad \text{(in centiliters)}$$

First assume that the standard deviation of the filling is known in advance. It is $\sigma = 1.5$ centiliters.

Perform the (one-sided; in which direction?) hypothesis test at the 5 % significance level. Formulate or calculate *explicitly*

- the model assumptions, H_0, H_A
- the rejection range
- the value of the test statistics and the test result
- the p value

Formulate the conclusion for the critical consumer in one sentence.

Problem 4.2

A bakery states that the rolls it produces have a minimum weight of 50 g with known standard deviation $\sigma = 3$ g. The weights are normally distributed.

A statistics student who is suspicious and suspects that the rolls are too light buys 16 rolls at the bakery and weighs all the rolls. He gets the following values (in g):

$$46,\ 48,\ 52,\ 49,\ 46,\ 51,\ 52,\ 47,\ 49,\ 44,\ 48,\ 51,\ 49,\ 50,\ 53,\ 47$$

($**$) a) Formulate the null and alternative hypothesis and carry out a hypothesis test at the 5 % significance level.

($**$) b) The student is concerned about the small sample size of 16 in his experiment. Therefore, he examines the weight of the rolls again, this time for 100 rolls. He gets the same average in the sample as for the 16 rolls in a).

Is the test decision the same as in a)? Justify your the answer.

Classical and Bayesian Statistic

Sample solution for Problems 4

Solution 4.1

Let X_i denote the content (in centiliters) of the ith wine bottle for $i = 1, \dots, n = 12$.

a) *Model*:

$$X_1, \dots, X_{12} \text{ i.i.d. } \sim \mathcal{N}(\mu, \sigma^2), \quad \sigma^2 = 1.5^2 \text{ known}$$

b) *Null hypothesis*:

$$H_0: \quad \mu = \mu_0 = 70$$

Alternative hypothesis:

$$H_A: \quad \mu < \mu_0$$

c) *Test statistics*:

$$\overline{X}_n$$

Distribution of test statistics under H_0 being true:

$$\overline{X}_n \sim \mathcal{N}\left(70, \frac{1.5^2}{12}\right)$$

d) *Significance level*:

$$\alpha = 5\,\%$$

e) *Rejection range for the test statistics*:

$$K = (-\infty, 69.29]$$

```
wine <- c(71, 69, 67, 68, 73, 72, 71, 71, 68, 72, 69, 72)

qnorm(p = 0.05, mean = 70, sd = 1.5 / sqrt(length(wine)))

[1] 69.28776
```

f) *Test decision*:

The observed mean is

$$\overline{x}_n = 70.25$$

```
mean(wine)

[1] 70.25
```

Thus $\overline{x}_n \notin K$ and H_0 is not rejected. It is therefore quite plausible that the wine merchant is bottling the wine correctly.

g) p-value

```
pnorm(q = 70.25, mean = 70, sd = 1.5 / sqrt(length(wine)))

[1] 0.7181486
```

The p-value of 0.718 is much higher than the significance level of 0.05 and therefore the test decision is, of course, the same as using the rejection range.

Note that the test was not really necessary, as the mean of the data is already greater than 70 and we are assuming that the mean is *less* than 70.

Solution 4.2

```
x <- c(46, 48, 52, 49, 46, 51, 52, 47, 49, 44, 48, 51, 49, 50, 53, 47)
```

a) Null hypothesis:

$$H_0 : \quad \mu = 50$$

Alternative hypothesis

$$H_A : \quad \mu < 50$$

```
pnorm(p = mean(x), mean = 50, sd = 3/sqrt(16))

[1] 0.0668072
```

The p-value of 0.067 is just above the significance level and thus the null hypothesis H_0 is *not* rejected. The average of the values *is* lower than 50 g, but not significantly lower.

b) The p-value becomes *very* much smaller

```
pnorm(p = mean(x), mean = 50, sd = 3/sqrt(100))

[1] 8.841729e-05
```

and is well below the significance level. In this case, the null hypothesis would be clearly rejected. The value 50 g is statistically significantly wrong.

The reason *why* the p-value becomes much smaller is that with an increasing number of measurements the certainty of where the true mean lies *increases*. Thus, for the mean 48.875 g is quite possible for a small number of rolls, provided that $\mu = 50$ g, the null hypothesis, is correct.

For a large number of measurements the *same* mean 48.875 g is already rather unlikely, always assuming that null hypothesis $\mu = 50$ g is correct.