

Homework 3

Pascal Pilz, k12111234

2023-10-24

Example 6

We have a sample of 11 persons, 2 of which smoke. We want to calculate two different 90% confidence intervals.

```
alpha <- 0.1
q <- 1 - (alpha/2)
n <- 11
k <- 2
p <- k/n
print(sprintf("p: %.3f, n*p*(1-p): %.3f", p, n*p*(1-p)))
```

```
## [1] "p: 0.182, n*p*(1-p): 1.636"
```

a) 90% asymptotic confidence interval for the true proportion of smokers

First we get the point estimator $\hat{\pi}$ given by $\hat{\pi} = p = \frac{k}{n}$, then we use the formula found in the slides to calculate the confidence interval for $\alpha = 0.1$:

```
val <- qnorm(q) * sqrt(p*(1-p)/n)
conf_lower <- p - val
conf_upper <- p + val
print(sprintf("P(true proportion in [%.3f,%.3f]) = %s",
              conf_lower, conf_upper, 1-alpha))
```

```
## [1] "P(true proportion in [-0.009,0.373]) = 0.9"
```

As we can see, the result is not very convincing.

b) 90% exact confidence interval for the true proportion of smokers

For this we use formula found on slide 14, which involves the F -distribution:

```
conf_lower <- k / (k + (n-k+1) * qf(q, 2*(n-k+1), 2*k))
conf_upper <- (k+1) * qf(q, 2*(k+1), 2*(n-k)) /
              ((n-k) + (k+1) * qf(q, 2*(k+1), 2*(n-k)))
print(sprintf("P(true proportion in [%.3f,%.3f]) = %s",
              conf_lower, conf_upper, 1-alpha))
```

```
## [1] "P(true proportion in [0.033,0.470]) = 0.9"
```

Example 7

We have the following number of bacteria per sample taken for 10 different sample:

```
x <- c(13,14,12,11,16,16,17,20,20,21)
n <- length(x)

sample_mean <- mean(x)
sample_std <- sd(x)

print(sprintf("mean %.3f, std %.3f", sample_mean, sample_std))

## [1] "mean 16.000, std 3.528"
```

a) 95% confidence interval for the true mean

```
alpha <- 0.05
t <- qt(1-(alpha/2), n-1)

val <- t * sample_std / sqrt(n)

conf_lower <- sample_mean - val
conf_upper <- sample_mean + val

print(sprintf("P(true mean in [%.3f,%.3f]) = %s",
              conf_lower, conf_upper, 1-alpha))

## [1] "P(true mean in [13.476,18.524]) = 0.95"
```

b) test if “true mean” differs from $\mu_0 = 14$

We want to construct a test that checks whether the “true mean” μ of bacteria count differs from $\mu_0 = 14$. We want to use a type I error of 5%, we can assume normality.

For this, let us define the following:

- Null hypothesis H_0 : $\mu = \mu_0$
- Alternative hypothesis H_1 : $\mu \neq \mu_0$

What we want to verify is that the true mean differs from 14, thus we want to verify the alternative hypothesis. For this, we show that the null hypothesis can be rejected.

In our example, a one-sample t-test is appropriate:

```
crit_val <- qt(1-(alpha/2), n-1)
test_stat <- abs(sqrt(n)*(sample_mean-14)/sample_std)

print(sprintf("critical value: %.3f, test statistic: %.3f",
              crit_val, test_stat))

## [1] "critical value: 2.262, test statistic: 1.793"
```

Since the critical value is greater than the test statistic, we cannot reject the null hypothesis.