Answer for ex3.2

Jiaqi Wang

2025-10-02

Question 1

1-1

That's because in this problem, we need to evaluate whether the MRD-negative complete remission rate at the end of the induction with ponatinib is significantly greater than the previously reported rate for second-generation tyrosine kinase inhibitors. That means testing whether the MRD negativity rate is greater than 16%. From a clinical perspective, the result of equivalence or inferiorty would not alter treatment, so we just need to test whether this treatment is a better one. Accordingly, a one-sided one-sample proportion test is the most appropriate statistical procedure, as we are only intererted in treatment improvement.

1-2

The MRD-negative complete remission rate at the end of induction for ponatinib was 43.0% (61/142).

1-3

```
prop.test(61, 142,
          p = 0.16,
          alternative = "greater",
          correct = FALSE)
```

```
##
##  1-sample proportions test without continuity correction
##
## data:  61 out of 142, null probability 0.16
## X-squared = 76.781, df = 1, p-value < 2.2e-16
## alternative hypothesis: true p is greater than 0.16
## 95 percent confidence interval:
##  0.3631946 1.0000000
## sample estimates:
##         p
## 0.4295775
```

Using a one-sided one-sample proportion test without continuity correction (H_0: $\pi \leq 0.16$), we obtain a one-sided p-value < 2.2e-16. We reject H_0 and conclude that the MRD-negativity rate (61/142 = 43.0%) is significantly greater than 16%, which means ponatinib is a better treatment than second-generation TKIs.

1-4

In this context, the one-sided p-value is the probability, under the null hypothesis that the true MRD-negativity rate is 16%, of observing a sample proportion as large as or larger than 61/142 (43.0%) .

Question 2

2-1

At end of induction, the MRD-negativity rate for ponatinib (among evaluable samples) was

$$x = 61 \text{ out of } n = 142, \text{ so } \hat{p} = x/n = 0.4296.$$

We test

$$H_0: p \leq 0.16 \text{ vs } H_A: p > 0.16.$$

Under H_0 the standard error is

$$SE_0 = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{0.16 \times 0.84}{142}} \approx 0.0305.$$

The z statistic is

$$z = \frac{\hat{p} - p_0}{SE_0} = \frac{0.4296 - 0.16}{0.0305} \approx 8.76.$$

The one-sided p-value is

$$p = \Pr(Z \geq 8.76) \approx 9.6 \times 10^{-19},$$

which is far below 0.05; thus we reject H_0.

These match the direction and magnitude of prop.test (extremely small p), confirming the function's output.

2-2

```r
# One-sided 95% lower confidence bound (since alternative="greater")
x  <- 61
n  <- 142
p0 <- 0.16

ci_one <- prop.test(x, n, p = p0, alternative = "greater", correct = FALSE)$c
onf.int
ci_one  # returns [lower, 1]
```

```
## [1] 0.3631946 1.0000000
## attr(,"conf.level")
## [1] 0.95
```

```
# Two-sided CI with conf.level = 0.90 → same lower bound as the 95% one-sided
 CI
ci_two <- prop.test(x, n, conf.level = 0.90, correct = FALSE)$conf.int
ci_two  # [lower, upper]; lower ≈ ci_one[1]
```

```
## [1] 0.3631946 0.4985937
## attr(,"conf.level")
## [1] 0.9
```

The one-sided 95% CI reports only a lower bound (about 0.363): with 95% confidence, the true MRD-negativity rate is at least 36%. Setting a two-sided CI to 90% reproduces the same lower bound (because 95% one-sided = 90% two-sided).

Question 3

First-generation drug (raw data available)

Advantages:

- Raw data allow detailed, patient-level analyses within the same trial.
- Baseline covariates can be adjusted, reducing confounding.
- Ensures consistency in outcome definitions and measurement methods.

Disadvantages:

- Sample size may be relatively small, limiting statistical power.
- Data come from one study population, reducing external generalizability.

Potential biases:

- Selection bias if the trial population is not representative of the broader patient population.

Second-generation drug (summary data from literature)

Advantages:

- Literature often summarizes larger or multicenter studies, improving external validity.
- Access is quick and inexpensive, since no raw data collection is needed.

Disadvantages:

- Lack of patient-level data prevents adjustment for confounding factors.

- Study definitions, eligibility criteria, or outcome measures may differ from the current trial.
- Publication bias may distort the available evidence.

Potential biases:

- Information bias due to inconsistent measurement.
- Temporal bias if historical controls differ in supportive care or diagnostic methods.
- Confounding bias because baseline differences cannot be adjusted.

Question 4

4-1

Choose a two-sided test, because we are interested in whether the imatinib MRD-negativity rate is different from 16% (the historical benchmark for second-generation TKIs). Both higher and lower values would matter in assessing whether imatinib is an appropriate control.

4-2

```
prop.test(15, 68, p = 0.16,
          alternative = "two.sided", correct = FALSE)

##
##  1-sample proportions test without continuity correction
##
## data:  15 out of 68, null probability 0.16
## X-squared = 1.8573, df = 1, p-value = 0.1729
## alternative hypothesis: true p is not equal to 0.16
## 95 percent confidence interval:
##  0.1384901 0.3325674
## sample estimates:
##         p
## 0.2205882
```

Interpretation: Since p ≈ 0.17 is much larger than 0.05, we do not reject H_0. The imatinib MRD-negativity rate (22.1%) is not significantly different from the historical 16%. This suggests that imatinib can be considered an appropriate control consistent with earlier-generation drugs.

Question 5

5-1

The total sample size for this test is n = 13, and the observed sample proportion is p = 6/13 .

5-2

```r
dbinom(0:13, 13, 0.25)           # probability distribution for X=0,…,13
```

```
##  [1] 2.375726e-02 1.029481e-01 2.058963e-01 2.516510e-01 2.097092e-01
##  [6] 1.258255e-01 5.592245e-02 1.864082e-02 4.660204e-03 8.630008e-04
## [11] 1.150668e-04 1.046062e-05 5.811453e-07 1.490116e-08
```

```r
sum(dbinom(0:13, 13, 0.25))      # should equal 1
```

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

The distribution gives the probabilities of 0–13 events when X Binomial(13,0.25). The sum of probabilities is 1, confirming it is a valid probability distribution.

5-3

```r
sum(dbinom(c(0, 6:13), 13, 0.25))
```

```
## [1] 0.1039699
```

By summing the probabilities of outcomes as or more extreme than 6, the two-sided p-value is ≈ 0.104.

5-4

```r
binom.test(6, 13, p = 0.25)
```

```
##
##  Exact binomial test
##
## data:  6 and 13
## number of successes = 6, number of trials = 13, p-value = 0.104
## alternative hypothesis: true probability of success is not equal to 0.25
## 95 percent confidence interval:
##  0.1922324 0.7486545
## sample estimates:
## probability of success
##              0.4615385
```

The exact binomial test gives a two-sided p-value = 0.104, a 95% confidence interval of approximately (0.192, 0.749), and a sample estimate of $\hat{p} = 0.462$. This matches the manual calculation above.