

Chapter 9 Exam Prep Worksheet:

Statypus Insight: The Bridge to Inference

We are finally crossing the bridge from pure probability into statistical inference. Instead of counting individual successes, we are now zooming out to track how the sample proportion (\hat{p}) behaves across thousands of different samples. The new standard error formula might look intimidating, but don't be fooled—the underlying engine is still the exact same Normal curve you have already mastered.

Raw Exam Question:

The historical survival rate for juvenile platypuses in the Snowy River is known to be 65% ($p = 0.65$). A researcher is monitoring a new cohort and collects a random sample of $n = 150$ juveniles. Let \hat{p} represent the proportion of juveniles in this sample that survive their first year. What is the mean value and the standard deviation of the sampling distribution of \hat{p} ?

Bill the Statypus says: Structural Rigor check! The exam is testing if you can find the “under-the-hood” curve. Don't confuse the population spread with the spread of the **sampling distribution**. As your sample size (n) gets larger, the sampling distribution gets skinnier!

Sally the Statypus says: Math check! The mean of your sampling distribution ($\mu_{\hat{p}}$) is always just the population proportion ($p = 0.65$). To find the Standard Deviation, we use the formula:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}.$$

Your Turn (Calculate the mean and standard deviation):

Statypus Insight: The Law of the Skinnier Curve

The Central Limit Theorem is the most powerful tool in the Statypus arsenal. It tells us that even if individual platypuses are unpredictable, the *average* behavior of a large group follows a predictable, Normal path.

Raw Exam Question & Output:

River regulation faces opposition on the basis of environmental issues. In a survey to assess public opinion regarding a proposed dam on the Snowy River, a random sample of $n = 800$ local residents included 536 who strongly opposed the dam. The local government wants to know if this sample provides convincing evidence that a majority of all local residents oppose the dam.

```
> binom.test(x = 536, n = 800, p = 0.5, alternative = "greater")
```

```
Exact binomial test
```

```
data: 536 and 800
```

```
number of successes = 536, number of trials = 800, p-value < 2.2e-16
```

```
alternative hypothesis: true probability of success is greater than 0.5
```

```
95 percent confidence interval:
```

```
0.6413203 1.0000000
```

```
sample estimates:
```

```
probability of success  
0.67
```

Exam Tasks:

- Set up the null and alternative hypotheses (H_0 and H_a) needed to test the claim.
- What is the value of the sample proportion (\hat{p}) for this study?
- Based on the R output, what is your conclusion at the $\alpha = 0.05$ significance level?
- Does the 95% confidence interval provided in the output support your decision? Explain.

Bill the Statypus says: Vocabulary trap! The word “majority” is a mathematical secret code. It means strictly greater than 50%. This tells you exactly what number to use for your null hypothesis ($p = 0.50$) and what symbol to use for your alternative hypothesis ($>$)!

Sally the Statypus says: Don’t be intimidated by $p\text{-value} < 2.2e-16$! That is just the computer’s way of saying the p -value is basically zero. It is essentially impossible to get 536 out of 800 people to oppose the dam if the true population was evenly split at 50/50.

Seneca the Statypus: The Engine and The Output

The manual formulas on Page 1 prove you understand how the machine works. The software on Page 2 proves you know how to drive it. Do not let the wall of text in the console distract you from the only two things that matter: your p -value and your confidence interval.

Your Turn (Record your answer on the blank portion on the top of the next page.)

Your Turn (Answers for Page 2):**Raw Exam Question 3:**

Using the same setup and R console output from the previous page, carefully interpret the 95% confidence interval in the context of the problem.

Bill the Statypus says: Fill in the blanks! Every confidence interval interpretation has the same skeleton structure: “We are [XX]% confident that the true [parameter] of [population context] is between [lower bound] and [upper bound].”

Sally the Statypus says: Statypus Standard Alert! Do NOT say “we are 95% confident that the *sample* proportion is between...” We already know the sample proportion is 0.67! The interval is our net to catch the true **population** proportion (p).

Your Turn (Write your formal interpretation in a complete sentence):