# Chapter 9 - Lecture 3
# Tests Concerning a Population Proportion

Yuan Huang

Mar 11, 2013

# Introduction

Let $p$ denote the proportion of individual or objects in a population who possess a specified property ("success"). Let $X$ denote the number of success within a sample of size $n$, then $X \sim N(n, p)$.

For doing the test about $p$, there are two cases:

1. When $n$ is small, the tests will be based on Binomial distribution.
2. When $n$ is large, the tests can be derived by applying the central limit theorem. Therefore, it's special case of approximate z tests.

# Large sample test for one sample proportion

- Null Hypothesis: $H_0 : p = p_0$
- Test statistic value: $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$
- Rejection Regions:
  - If $H_A : p > p_0$, $z \geq z_\alpha$
  - If $H_A : p < p_0$, $z \leq -z_\alpha$
  - If $H_A : p \neq p_0$, $z \leq -z_{\alpha/2}$ and $z \geq z_{\alpha/2}$
- There are two conditions that we need to check, so that the use of this test is valid. We need:
  - $np_0 \geq 10$
  - $n(1 - p_0) \geq 10$

# Problem 9.38

A random sample of 150 recent donations at a blood bank reveals that 82 were type A blood. Does this suggest that the actual percentage of type A donations differs from 40%. Carry out a test of the appropriate hypotheses using a significance level of 0.01.

# Problem 9.38(cont.)

Solution: first check the conditions:
$np_0 = 150 \times 0.4 = 60, n(1 - p_0) = 150 \times 0.6 = 90$

Step 1: The hypotheses: $H_0 : p = 0.4$ vs $H_1 : p \neq 0.4$, where $p$ is proportion of type A donations in the population.

Step 2: Test statistics value $z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{\frac{82}{150} - 0.4}{\sqrt{0.4(1-0.4)/n}} = 3.67$

Step 3: The corresponding null distribution for $Z$ is $N(0,1)$.

Step 4 Select the significant level $\alpha$, The rejection region is of the form $z \leq -z_{\alpha/2}$ and $z \geq z_{\alpha/2}$ where for $\alpha = 0.01$, $z_{\alpha/2} = 2.58$.

step 5: Since $3.67 > 2.58$, we have significant evidence to reject $H_0$ and say that the percentage of type A donations is not 40% under significance level 0.01

# Calculating Type II error (given true $p = p'$)

1. If $H_A : p > p_0$

$$\beta(p') = P\left(z < \frac{p_0 - p' + z_\alpha \sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$$

2. If $H_A : p < p_0$

$$\beta(p') = 1 - P\left(z < \frac{p_0 - p' - z_\alpha \sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)$$

3. If $H_A : p \neq p_0$,

$$\begin{aligned}\beta(p') &= P\left(z < \frac{p_0 - p' + z_{\alpha/2} \sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right) \\ &\quad - P\left(z < \frac{p_0 - p' - z_{\alpha/2} \sqrt{p_0(1-p_0)/n}}{\sqrt{p'(1-p')/n}}\right)\end{aligned}$$

# Small sample test for one sample proportion

The procedure when the sample size $n$ is small are based directly on the binomial distribution rather than the normal approximation. In the small sample case, consider only one-sided tests.

- The book gives the case when the alternative is $H_\alpha : p > p_0$. Therefore, I will give detail procedures here for alternative $H_\alpha : p < p_0$.
- Review of Binomial Distribution

  1. If $X \sim \text{Binomial}(n, p)$, $E(X) = np$ and $V(X) = np(1-p)$
  2. $P(X = x | X \sim \text{Binomial}(n, p)) = \binom{n}{x} p^x (1-p)^{(n-x)}$
  3. Notation: $B(x, n, p) \doteq P(X \leq x | X \sim \text{Binomial}(n, p))$

# one-sided test for one sample proportion with $H_\alpha : p < p_0$

Step 1: The hypotheses: $H_0 : p = p_0$ vs $H_1 : p < p_0$.

Step 2: Test statistics $X$, the number of success events.

Step 3: The corresponding null distribution for $X$. Under $H_0$, $X \sim \text{Binomial}(n, p_0)$.

Step 4 Select the significant level $\alpha$, The rejection region is of the form $\{x : x \leq c\}$ . The critical value $c$ satisfies that $B(c, n, p_0) < \alpha$ and $B((c + 1), n, p_0) > \alpha$. (Hence this is not exact $\alpha$ test.) Hence $c = 15$ and the rejection region is $\{x \leq 15\}$.

step 5: Based on the observation $x$ to make the decision.

# one-sided test for one sample proportion with $H_\alpha : p < p_0$ (cont.)

## Type II error

If the true value of $p$ is $p' < p_0$, then the type II error is calculated as

$$\begin{aligned}\beta(p') &= P[H_0 \text{is not rejected when } X \sim \text{Binomial}(n, p')] \\ &= P[X \geq (c+1) \text{when } X \sim \text{Binomial}(n, p')] \\ &= 1 - B(c; n, p')\end{aligned}$$

Note: R command *pbinom(x,n,p)*.

# Example 9.13

A plastic manufacturer has developed a new type of plastic trash can and proposes to sell them with an unconditional 6-year warranty. To see weather this is economically feasible, 20 prototype cans are subjected to an accelerated life test to simulate 6 years of use. The proposed warranty will be modified only if the sample data strongly suggests that fewer than 90% of such cans would survive the 6-year period.

# Example 9.13 (cont.)

Step 1: The hypotheses: $H_0 : p = 0.9$ vs $H_a : p < 0.9$.

Step 2: Test statistics $X$, the number among the 20 that survive.

Step 3: The corresponding null distribution for $X$. Under $H_0$, $X \sim \text{Binomial}(20, 0.9)$.

Step 4 Let the significant level $\alpha = 0.05$, Then the critical value $c$ satisfies that $B(c, 20, 0.9) < 0.05$.
$B(c, n, p) \doteq P(X \leq c | X \sim \text{Binomial}(n, p))$. Given

1. B(15, 20,0.9) = 0.043
2. B(16, 20,0.9) = 0.133

Hence $c = 15$ and the rejection region is $\{x \leq 15\}$.

step 5: If $x = 14$, then $H_0$ would be rejected, and the data favors $p < 0.9$ at the significance level 0.05.

# Example 9.13 (cont.)

**Type II error**

If the true value of $p$ is $p' = 0.8$, then the type II error is calculated as

$$\begin{aligned}\beta(0.8) &= P[H_0 \text{is not rejected when } X \sim \text{Binomial}(20, 0.8)] \\ &= P[X \geq 16 \text{when } X \sim \text{Binomial}(20, 0.8)] \\ &= 1 - B(15; 20, 0.8) = 1 - 0.37 = 0.63\end{aligned}$$

Note: R command *pbinom(15,20,0.8)*.