Hypothesis Testing for Variance and Proportion

October 3, 2013

Outline

- **Hypothesis Testing for Variance**
- **Hypothesis Testing for Proportion**
- **Sample Size and Power of a Test**

Test for Variances

There are three forms for tests of hypotheses on the variance of a distribution:

<u>Right Tailed test</u>

$$H_0 : \sigma^2 = \sigma_o^2, \qquad H_1 : \sigma^2 > \sigma_o^2$$

<u>Left Tailed test</u>

$$H_0 : \sigma^2 = \sigma_o^2, \qquad H_1 : \sigma^2 < \sigma_o^2$$

<u>Two Tailed test</u>

$$H_0 : \sigma^2 = \sigma_o^2, \qquad H_1 : \sigma^2 \neq \sigma_o^2$$

The test statistic used to test each of these is

$$\chi_o^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

When sampling from a normal distribution, this statistic is known to follow a chi-square distribution with $n-1$ degrees of freedom under the null hypothesis.

Hypothesis Testing for Variance

Decision criteria:

Hypothesis	Rejection Criteria
$H_o : \sigma^2 = \sigma_o^2 \quad Vs. \quad H_1 : \sigma^2 > \sigma_0^2$	$\chi_0^2 > \chi_{\alpha,n-1}^2$
$H_o : \sigma^2 = \sigma_o^2 \quad Vs. \quad H_1 : \sigma^2 < \sigma_0^2$	$\chi_0^2 < \chi_{1-\alpha,n-1}^2$
$H_o : \sigma^2 = \sigma_o^2 \quad Vs. \quad H_1 : \sigma^2 \neq \sigma_0^2$	$\chi_0^2 < \chi_{1-\alpha/2,n-1}^2$ Or $\chi_0^2 > \chi_{\alpha/2,n-1}^2$

Example

The quality control office of a large hardware manufacturer received a complaint about the diameter variability of its 4 cm washers. A random sample of 20 washers are chosen and the diameters are recorded as below:
4.06 4.02 4.04 4.04 3.97 3.87 4.03 3.85 3.91 3.98
3.96 3.90 3.95 4.11 4.00 4.12 4.00 3.98 3.92 4.02
Conduct an appropriate hypothesis test at $\alpha = 0.05$ whether the variability is greater than $0.004 cm^2$
Solution We would like to test

$$H_0 : \sigma^2 = 0.004$$
$$H_1 : \sigma^2 > 0.004$$

For the subject that we have

$$\begin{aligned} s^2 &= 0.005318684 \\ \chi_0^2 &= 19 * s^2/0.004 = 25.26375 \\ p - value &= 1 - pchisq(25.26375, 19) = 0.1520425 \end{aligned}$$

Example

We can use R to perform the hypothesis testing as follows:

```
> library(TeachingDemos) # Install TeachingDemos package
> x
 [1] 4.06 4.02 4.04 4.04 3.97 3.87 4.03 3.85 3.91 3.98
 [11]3.96 3.90 3.95 4.11 4.00 4.12 4.00 3.98 3.92 4.02
> sigma.test(x,sigmasq=0.004, alternative="greater")

        One sample Chi-squared test for variance
        data:  x
X-squared = 25.2637, df = 19, p-value = 0.152
alternative hypothesis: true variance is greater than 0.004
95 percent confidence interval:
 0.003352461         Inf
sample estimates:
   var of x
0.005318684
```

Decision: Since p-value=0.152 $>$ 0.05 we fail to reject the null hypothesis and conclude that there is insufficient evidence to suggest that the variance has increased from 0.004 cm^2

Testing for equality of variances when sampling from independent Normal distributions

Let us consider a problem of comparing the variances of two normal population. Consider independent random sample of size n_1 and n_2 are taken from two populations and we are interested in testing the hypothesis

$$\begin{aligned} H_0 &: \sigma_1^2 = \sigma_2^2 \\ H_1 &: \sigma_1^2 \neq \sigma_2^2 \end{aligned}$$

which is equivalent to

$$\begin{aligned} H_0 &: \frac{\sigma_1^2}{\sigma_2^2} = 1 \\ H_1 &: \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \end{aligned}$$

We know that the test statistic

$$F_0 = \frac{s_1^2}{s_2^2}$$

has F-distribution with $n_1 - 1$ numerator degrees of freedom and $n_2 - 1$

Testing for equality of variances when sampling from independent Normal distributions

<u>Decision Criteria:</u>

Hypothesis	Rejection Criteria
$H_o : \sigma_1^2 = \sigma_2^2 \quad Vs. \quad H_1 : \sigma_1^2 > \sigma_2^2$	$F_0 > F_{\alpha, n_1 - 1, n_2 - 1}$
$H_o : \sigma_1^2 = \sigma_2^2 \quad Vs. \quad H_1 : \sigma_1^2 < \sigma_2^2$	$F_0 < F_{1-\alpha, n_1 - 1, n_2 - 1}$
$H_o : \sigma_1^2 = \sigma_2^2 \quad Vs. \quad H_1 : \sigma_1^2 \neq \sigma_2^2$	$F_0 > F_{\alpha/2, n_1 - 1, n_2 - 1}$ or $F_0 < F_{1-\alpha/2, n_1 - 1, n_2 - 1}$

Example

At a 0.05 level of significance, test whether the following two data have different variances

x: 60, 39, 55, 58, 63, 45, 50

y: 42, 38, 25, 33, 51, 37, 40

```
> x=c(60, 39, 55, 58, 63, 45, 50)
> y=c(42, 38, 25, 33, 51, 37, 40)
> var.test(x,y)
        F test to compare two variances
data:  x and y
F = 1.1637, num df = 6, denom df = 6, p-value = 0.8587
alternative hypothesis:true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1999552 6.7723953
sample estimates:
ratio of variances
           1.16369
```

Since the p-value=0.8587 $>$ 0.05, the null hypothesis of equal variances is not rejected.

Example

A manufacturer of lithium batteries has two production facilities, A and B. Fifty randomly selected batteries with an advertised life of 180 hours are selected, and tested. The lifetimes are stored in (facilityA). Fifty randomly selected batteries with an advertised life of 200 hours are selected, and tested. The lifetimes are stored in (facilityB). Test for the equality of variance.

```
> library(PASWR)
> data(Battery)
> attach(Battery)
> var.test( facilityA, facilityB)
        F test to compare two variances
data:  facilityA and facilityB
F = 0.5766, num df = 49, denom df = 49, p-value = 0.05672
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.3272049 1.0160721
sample estimates:ratio of variances
                   0.5765967
```

Since $0.05672 > 0.05$, the null hypothesis of equal variances is not rejected.

Testing for Proportion

We know that the maximum likelihood estimate of the population proportion p is the sample proportion $\hat{p}$. The asymptomatic properties of MLE allows that

$$\hat{p} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

Therefore, to test

$$H_0 : p = p_0$$
$$H_a : p \neq p_0$$

The test statistics is given by

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

as long as $np_0 \geq 10$ and $n(1-p_0) \geq 10$. (So that the normal approximation for the binomial distribution works.)

Hypothesis Testing for Proportion

Decision criteria:

Hypothesis	Rejection Criteria
$H_o : p = p_o \quad Vs. \quad H_a : p > p_0$	$Z > Z_\alpha$
$H_o : p = p_o \quad Vs. \quad H_a : p < p_0$	$Z < -Z_\alpha$
$H_o : p = p_o \quad Vs. \quad H_a : p \neq p_0$	$\lvert Z \rvert > Z_{\alpha/2}$

Example

An airline claims that, on average, 5% of its flights are delayed each day. On a given day, of 500 flights, 50 are delayed. Test the hypothesis that the average proportion of delayed flights is 5% at the 0.01 level.
We want to test

$$H_0 : p = 0.05$$
$$H_a : p > 0.05$$

The test statistic is

$$\frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{\frac{50}{500} - 0.05}{\sqrt{0.05(1-0.05)/500}} = 5.13$$

`p-value= 1-pnorm(5.13)=0.00000014487.`
Decision: Reject H_0

Example

An airline claims that, on average, 5% of its flights are delayed each day. On a given day, of 500 flights, 50 are delayed. Test the hypothesis that the average proportion of delayed flights is 5% at the 0.01 level.
We want to test

$$H_0 : p = 0.05$$
$$H_a : p > 0.05$$

```
> binom.test(50,500,alternative="greater",p=0.05)
        Exact binomial test
data:  50 and 500
number of successes = 50, number of trials = 500, p-value = 3.576e-06
alternative hypothesis:true probability of success is greater than 0.05
95 percent confidence interval:
 0.07873857 1.00000000
sample estimates:
probability of success
                   0.1
```

With and Without Continuity Correction

```
> prop.test(50,500, alt="greater", p=0.05, correct=TRUE)
        1-sample proportions test with continuity correction
data:  50 out of 500, null probability 0.05
X-squared = 25.2737, df = 1, p-value = 2.487e-07
alternative hypothesis: true p is greater than 0.05
95 percent confidence interval:
 0.07914173 1.00000000
sample estimates:
  p
0.1

> prop.test(50,500, alt="greater", p=0.05, correct=FALSE)
        1-sample proportions test without continuity correction
data:  50 out of 500, null probability 0.05
X-squared = 26.3158, df = 1, p-value = 1.45e-07
alternative hypothesis: true p is greater than 0.05
95 percent confidence interval:
 0.08003919 1.00000000
sample estimates:
  p
0.1
```

Example

An insurance company states that 90% of its claims are settled within 30 days. A consumer group selected a simple random sample of 75 of the companys claims to test this statement. The consumer group found that 55 of the claims were settled within 30 days. At the 0:05 significance level, test the companys claim that 90% of its claims are settled within 30 days.
We want to test

$$H_0 : p = 0.9$$
$$H_a : p < 0.9$$

The Power of a test

Power is one of the most important things in experimental design. When a hypothesis test does not reject the null hypothesis when its false a type II error has been made. The power of the test is the probability of rejecting the null hypothesis when it is false, in other words that the test will not make a type II error. Thus power is defined as

$$\textit{Power} = 1 - P(\text{Type II Error}) = P(\text{Reject } H_0 \text{ when } H_0 \text{ is false})$$

Since power is the probability of correctly rejecting the null hypothesis when it is false it makes sense that we would like this as large as possible. To do this, as shown in the formula above, we would like the probability of a type II error to be as small as possible.
There are two different aspects of power analysis.

- Calculate the necessary sample size for a specified power
- Calculate the power when given a specific sample size

The Power of a test

The power of the test is directly related to the number of individuals per group (n) , the amplitude of the differences we want to detect(also known as effect size), within group variability(σ) and the type I error (α). For t-tests, use the power.t.test function in R

```
power.t.test(n = NULL, delta = NULL, sd = 1, sig.level = 0.05,
             power = NULL,
             type = c("two.sample", "one.sample", "paired"),
             alternative = c("two.sided", "one.sided"),
             strict = FALSE)
```

Note that for a single sample test

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_a - \mu_o)^2}$$

In the R code $\delta = \frac{\mu_a - \mu_o}{\sigma}$

Example

A company that manufactures light bulbs claims that a particular type of light bulb will last 850 hours on average with standard deviation of 50. A consumer protection group thinks that the manufacturer has overestimated the lifespan of their light bulbs by about 40 hours. How many light bulbs does the consumer protection group have to test in order to prove their point with reasonable confidence?

```
> power.t.test(d=(850-810)/50,power=0.9,sig.level=0.05,
type="one.sample",alternative="two.sided")

     One-sample t test power calculation

              n = 18.44623
          delta = 0.8
             sd = 1
      sig.level = 0.05
          power = 0.9
    alternative = two.sided
```