Onkar Kurkute (571-276-3474)

Aim:
To find Statistically Difference between samples of two coins with different probability of face values

Abstract:
Two coins, one with 5% probability and other with 2% probability for face value of 1 are sampled over 10000 tosses. To find Statistical Difference, we have conducted 3 types of test for hypothesis testing, additionally we have found that if sample size of 5000 is not enough to prove the Statistical Difference it should be increased to 10000 tosses. As a result of those tests we have found that, those two coins are Statistically Different.

Introduction:
Firstly, we implemented a program that will sample data for these two coins. Data reading per 100 tosses are stored in an array, so totally we have 100 readings (of 100 tosses) for each of the coins. Implementation of program enables independent sampling for these coins. These reading are undergone the hypothesis testing methods such as f-test, t-test, z-test and chi square test, each prove the various statistical differences between these coin samples. Sample size is calculated as per sample size formula with respect to error rate and standard deviation.

Experimental Methods:
The program implementation uses two objects as two different coins and it will produce the random face value (0 or 1) with give probability (2% or 5%) using Math.random() function in Java as follows:

```
Boolean Face-value= (Math.random())< =Probability/100;
```

This call to generates random face value whose probability of being true is given probability.

Generated face values are counted for each 100 tosses as coin with 2% probability will have 2 true responses and one with 5% probability will have 5 true response which are both non- fractional and such 100 readings are added to tosses[] array.

We used calculateMeasures(sampleSize) method where we calculate all the measures such as sum, average, variance and standard deviation for each coin.

Also, tests(coin1, coin2, sampleSize) will calculate the values of F statistic, t- statistic, z- statistic and $X^2$ statistic for f-test, t-test, z-test and chi square test respectively with help of measures calculated earlier.

The similar tests are done in MS-Excel and compared the results. In both scenarios test cases yield into same result.

Results:

**To determine the sample size,**

We have selected confidence level of 95% and confidence interval of 5% and with multiples runs of sampling for 5% coin, Standard Deviation is at max 2.5 which is > Standard Deviation of 2% coin which is at max 1.7

So, $\mu_{5\%} = 5$ and CI= .05, $Z_{95\%}=1.96$, $\sigma_{5\%} = 2.5$

As CI= $z^* \sigma/n^{1/2}$ where n is sample size.

Therefore, sample size $n = z^{2*} \sigma^2/CI^2$

$n = 1.96^{2*}2.5^2/.05^2 = 9604$

This proves that sample size of 5000 is not enough to prove the statistical significance between two coins as it is less than 9604. So sample size should be 10000 tosses.

**To calculate statistical measures, we used following formulae**

$$\text{Mean } \mu = \frac{1}{n}\sum x_i$$

$$\text{Variance } \sigma^2 = \frac{1}{n}\sum (x_i - \overline{x})^2$$

$$\text{Standard Deviation } \sigma = \sqrt{\frac{1}{n}\sum (x_i - \overline{x})^2}$$

To determine Statistical Difference,

*We have done all the tests in MS-Excel as well as implemented in program to calculate the f, t, z and chi statistic. A MS-Excel sheet is attached with this report which shows the calculations and critical values of those statistics to be compared.*

**Test 1- F test**

F-test is used for comparison of two variances

Let $X_1, ..., X_n$ and $Y_1, ..., Y_m$ be independent and identically distributed samples from two coin's populations which each have a normal distribution.

The expected values for the two populations are different, and the hypothesis to be tested is that the variances of both coins are equal (null hypothesis).

Let sample means be the,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } \overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i$$

Let sample variances be the,

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \text{ and } S_Y^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(Y_i - \overline{Y}\right)^2$$

Then the F-test statistic is,

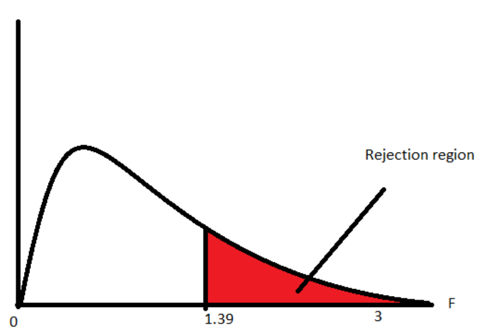
$$F = \frac{S_X^2}{S_Y^2}$$ ……….where $S^2_x > S^2_y$

It has an F-distribution with $n - 1$ and $m - 1$ degrees of freedom (here n and m are equal).

If the null hypothesis of equality of variances is true. Otherwise they differ with respect to distribution. The null hypothesis is rejected if $F$ is either too large or too small. This test allows us to select equal or unequal variance for t-test.

**Null hypothesis**: $H_0$ : Variances of both coins are equal
**Alternate hypothesis**: $H_a$ : Variances of both coins are different
As a result of calculation on various samples for both coins, $F_{c1,c2} > F_{Critical(1.394)}$

Which shows that variances of both coins are different variances.




Area in red color shows the rejection region.

**Test 2- z test**

z test is used here to find difference between means two samples

These are two normally distributed but independent populations, Standard Deviation σ is known

**Hypothesis test**

**Formula**: $$z = \frac{\bar{x}_1 - \bar{x}_2 - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the means of the two samples, Δ is the hypothesized difference between the population means (0 if testing for equal means), $\sigma_1$ and $\sigma_2$ are the standard deviations of the two populations, and $n_1$ and $n_2$ are the sizes of the two samples.

**Null hypothesis**: $H_0$: Means of both coins have a hypothesized difference of equal to or more than zero or
$H_0: \mu_1 - \mu_2 \geq 0$

**Alternate hypothesis**: $H_a$ : Means of both coins have a hypothesized difference of less than 0 or
$H_a : \mu_1 - \mu_2 < 0$
But with respect to this case means are different so we have new hypothesis as,

**One Tail Test:** $H_0: \mu_1 \geq \mu_2$, $H_a: \mu_1 < \mu_2$

**Two Tail Test:** $H_0: \mu_1 = \mu_2$, $H_a: \mu_1 \neq \mu_2$

In practice, the two-sample *z*-test is not used often, because the two population standard deviations $\sigma_1$ and $\sigma_2$ are usually unknown. Instead, sample standard deviations and the *t*-distribution are used.

As a result of calculation on various samples for both coins, $z_{c1,c2} < z_{Critical}$ one tailed(1.65)

Which shows, $\mu$ -$coin_1 \geq \mu$ -$coin_2$ i.e. means both coins are different and mean of coin1 is greater than coin2, this also concludes as Means of both coins have a hypothesized difference of is more than zero.

Area after the colored lines shows the rejection region for sub-sequent tests.

Test 3- t Test

This test is used only when the two population variances are not equal, hence must be estimated separately. The *t* statistic to test whether the population means are different is calculated as:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{s_{\overline{X}_1 - \overline{X}_2}}$$

where

$$s_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Here *S* is the unbiased estimator or standard error of the variance of the two samples, $n_i$ = number of participants in group *i*, *i*=1 or 2. In this case $n_1 = n_2$

**Null hypothesis**: $H_0$: Means of both coins have a hypothesized difference of equal to or more than zero or

$H_0$: $\mu_1 - \mu_2 \geq 0$

**Alternate hypothesis**: $H_a$ : Means of both coins have a hypothesized difference of less than 0 or

$H_a$ : $\mu_1 - \mu_2 < 0$

But with respect to this case means are different so we have new hypothesis as,

**One Tail Test:** $H_0$: $\mu_1 \geq \mu_2$, $H_a$: $\mu_1 < \mu_2$

**Two Tail Test:** $H_0$: $\mu_1 = \mu_2$, $H_a$: $\mu_1 \neq \mu_2$

As a result of calculation on various samples for both coins, $t_{c1,c2} < t_{Critical}$ one tailed(1.65)

Which shows, $\mu$ -coin$_1 \geq \mu$ -coin$_2$ i.e. means both coins are different and mean of coin1 is greater than coin2, this also concludes as Means of both coins have a hypothesized difference of is more than zero.

Area after the colored lines shows the rejection region for sub-sequent tests.

**Test 3- chi-square Test**

A chi-square goodness of fit test allows us to test whether the observed proportions for a categorical variable differ from hypothesized proportions. And it is calculated as,

$$\chi^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i}$$

where $o_i$ are observed values and $e_i$ are expected values

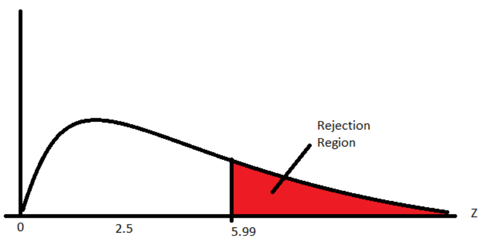
$H_0$: Samples of Coin1 and Coin2 are independent

$H_a$: Samples of Coin1 and Coin2 are not independent

With respect to two coins, we have degree of freedom=2 so $X^2_{critical}=5.99$ *with ME=.05*

As a result of calculation on various samples for both coins, $X^2_{c1,c2} < X^2_{Critical(5.99)}$

Which shows that samples of Coin1 and Coin2 are independent.




Area in red color shows the rejection region

Conclusion

As a result of sample size calculation with respect to Margin of Error(ME) proved that or ME=.05. We should use more than 9604 samples to prove some statistical significance with 95% of confidence interval, we have used 10000 tosses for all the tests.

As a result of F-test we found that both coins have different variances i.e. they have different distribution with respect to each other. This also helps us to use unequal variances for t- test, which proved that $\mu\text{-coin}_1 \geq \mu\text{-coin}_2$ i.e. mean for sample of coin with 5% probability > mean for sample of coin with 2% probability. In Chi-square test proved that these two coin samples are different and independent to each other.

Finally, this proves that these two coins are totally different with respect to statistically significance based on difference on variances, mean and goodness of fit.

References:

Wikipedia

StatTrek.com

YouTube.com