Harold's Statistics
Hypothesis Testing
Cheat Sheet

23 June 2022

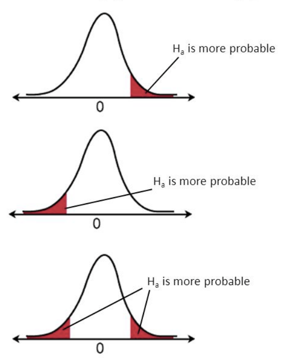
Hypothesis Terms	Definitions
Significance Level (α)	Defines the strength of evidence in probabilistic terms. Specifically, alpha represents the probability that tests will produce statistically significant results when the null hypothesis is correct. In most fields, $\alpha = 0.05$ is used most often.
Confidence Level (c)	The percentage of all possible samples that can be expected to include the true population parameter. $\alpha + c = 1$
Confidence Interval	A range of values within which you are fairly confident that the true value for the population lies. (e.g., 69% ± 3.8%)
Critical Value (z^*)	z^* is the critical value of a standard normal distribution under H_0. Critical values divide the rejection and non-rejection regions. Set using p-values or to a threshold value of 0.05 (5%) or 0.01 (1%), but always ≤ 0.10 (10%).
Test Statistic (z_{data})	A value calculated from sample data during hypothesis testing that measures the degree of agreement between the sample data and the null hypothesis. If z_{data} is inside the rejection region, demarked by z^*, then we can reject the null hypothesis, H_0.
p-value	Probability of obtaining a sample "more extreme" than the ones observed in your data, assuming H_0 is true.
Hypothesis	A premise or claim that we want to test.
Null Hypothesis: H_0	Currently accepted value for a parameter (middle of the distribution). Is assumed true for the purpose of carrying out the hypothesis test. • Always contains an "=" {=, ≤, ≥} • The null value implies a specific sampling distribution for the test statistic • H_0 is the middle of the normal distribution curve at $z = 0$. • Can be rejected, or not rejected, **but NEVER supported**
Alternative Hypotheses: H_a	Also called Research Hypothesis or H_1. Is the opposite of H_0 and involves the claim to be tested. Is supported only by carrying out the test if the null hypothesis can be rejected. • Always contains ">" (right-tailed), "<" (left-tailed), or "≠" (two-tailed) [tail selection is important] • Can be supported (by rejecting the null), or not supported (by failing or rejecting the null), **but NEVER rejected**

Hypothesis Testing	Steps
Hypothesis Testing (for one population)	1) Claim: Formulate the **null** (***H_0***) and the **alternative** (***H_a***) hypothesis 2) Graph: Sketch and label critical value (left-tailed, right-tailed, two-tailed) 3) Decision Rule: Use significance level (α), confidence level (c), confidence Interval, or critical value z*. e.g., We will reject H_0 if $z_{data} > 1.645$. 4) Critical Value: Determine **critical values** (z*) to mark the rejection regions 5) Test Statistic: Calculate the **test statistic** (z_{data} or t_{data}) from the sample data 6) p-Value: Use the test statistic to find the p-value 7) Conclusion: Reject the null hypothesis (supporting the alternative hypothesis) otherwise fail to reject the null hypothesis, then state claim

1) Claim: Formulate Hypothesis

If claim consists of …	*then the hypothesis test is*	*and is represented by…*
"is equal to", "is exactly", "is the same as", "is between" "is at least" "is at most"	Two-tailed = Left-tailed ≤ Right-tailed ≥	H_0
"is not equal to", "is different from", "has changed from" "is less than", "is below", "is lower or smaller than", "reducing" "is greater than", "is above", "is longer or bigger than"	Two-tailed ≠ Left-tailed < Right-tailed >	H_a
Make sure $H_0 + H_a$ = all possible outcomes.		

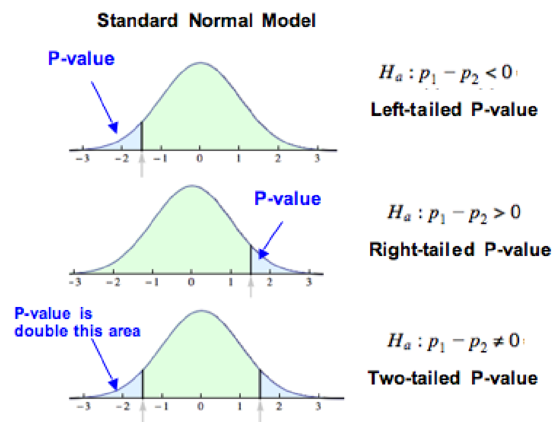
Types of Hypothesis Tests



Right-tail test
Ha: $\mu >$ value

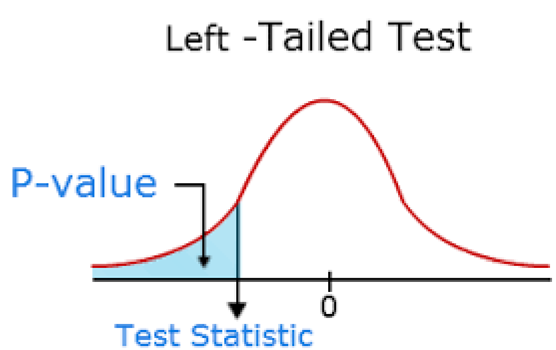
Left-tail test
Ha: $\mu <$ value

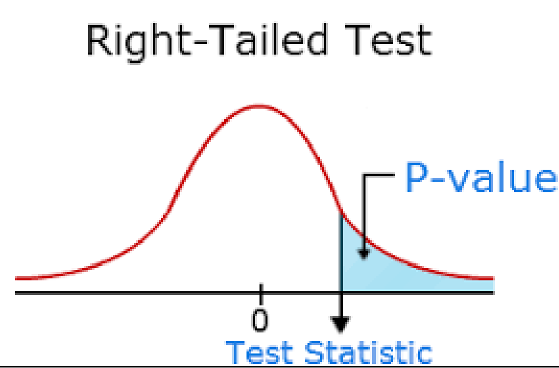
Two-tail test
Ha: $\mu \neq$ value

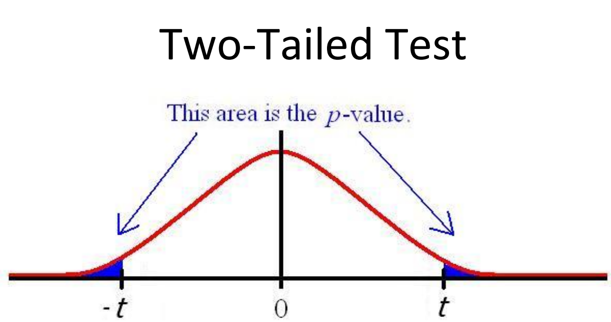


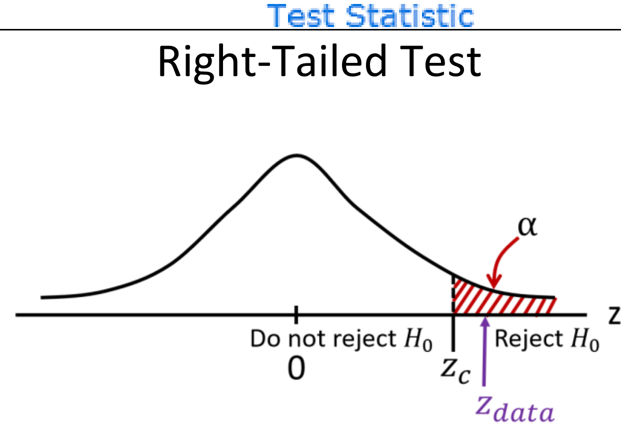
2) Graph: Sketch and Label

Sketch and label critical value (z^* or z_c).	Look at the direction of the inequality symbol in H_a to determine where to shade.









3) Decision Rule	
p-value	Use probability value to determine z_c in a Normal distribution table.
Significance level (α)	$\alpha = 1 - c$ Usually at a threshold value of 0.05 (5%) or 0.01 (1%), but always ≤ 0.10 (10%). The significance level α is the area under the curve outside the confidence interval.
Confidence Level (c)	$c = 1 - \alpha$ With a confidence of 0.95 (95%) or 0.99 (99%), but always ≥ 0.90 (90%).
Confidence Interval for μ	A 95% confidence interval means that the interval calculated has a probability of 95% containing the population mean, μ. σ known, normal population or large sample (n) $z\ interval = \bar{x} \pm SE(\bar{x})$ $= \bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$ $= [\bar{x} - SE(\bar{x}), \bar{x} + SE(\bar{x})]$ $\frac{\alpha}{2} = \frac{1-c}{2}$ $z^* = z_{\alpha/2} = z - score\ for\ probabilities\ of\ {}^{\alpha}/_{2}\ (two - tailed)$
Examples	We will reject the null hypothesis (H_0) if: • Significance level (α) is less than 5% • Confidence level (c) is greater than 95% • Confidence interval is between 5% and 95% (± 5%) `zdata > z* in a right-tailed test`

Python	`import scipy.stats as st`
	`n = 100`
	`df = n - 1`
	`mean = 219`
	`stderr = (sd = 35.0)/(n ** 0.5)`
	`print(st.t.interval(0.95, df, mean, stderr))`

4) Determine Critical Values (z*) / Rejection Region	
Critical Values (z*)	Determine z* by looking up α, c, or p-values in a standard normal distribution table. Two-tailed tests have two values for z*.

Significance Level (α)	Confidence Level (c)	Critical Value
α = 0.10	c = 0.90	z* = 1.645
α = 0.05	c = 0.95	z* = 1.960
α = 0.01	c = 0.99	z* = 2.576

5) Calculate Test Statistic (z_{data}) or z-score

Population Mean (μ) / Sample Mean ($\bar{x}$)	$z_{data} = \frac{\bar{x}-\mu}{SE(x)} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$	Variance known. Assumes data is normally distributed **or** $n \geq 30$ since t approaches standard normal Z if n is sufficiently large due to the CLT.
	$t_{data} = \frac{\bar{x}-\mu}{SE(\bar{x})} = \frac{\bar{x}-\mu}{s/\sqrt{n}}$	Variance unknown. t distribution, $df = n-1$ under H_0.
Population Proportion (p) / Sample Proportion ($\hat{p}$)	$z_{data} = \frac{\hat{p}-p}{SE(\hat{p})} = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$	Population proportion known. To be statistically significant, this assumes $np \geq 15$ $\boldsymbol{and}$ $n(1-p) \geq 15$.
	Worst Case: $p = 0.50$ $z_{data} = \frac{\hat{p}-p}{SE(\hat{p})} = (2\hat{p}-1)\sqrt{n}$	Population proportion unknown.

Python (1 mean)

```
from statsmodels.stats.weightstats import ztest
import pandas as pd
from statsmodels.stats.proportion import
proportions_ztest
scores = pd.read_csv('http://data-
analytics.zybooks.com/ExamScores.csv')
print(ztest(x1 = scores['Exam1'], H0_value = 86))
print(st.ttest_1samp(scores['Exam1'], H0_value = 82))
print(proportions_ztest(count, nobs, value, prop_var
= value))
```

```
(-2.5113146627890988,
0.012028242796839027)
```

z-score = 2.511
p-value = 0.0120 / 2 = 0.0060

```
Ttest_1sampResult(statist
ic=0.5327, pvalue=0.5966)
```

Python (2 means)

```
from statsmodels.stats.weightstats import ztest
sample1 = [21, 28, 40, 55, 58, 60]
sample2 = [13, 29, 50, 55, 71, 90]
print(ztest(x1 = sample1, x2 = sample2, value = 0))
```

```
(-0.5801720810890816 9,
0.5617985790046 4247)
```

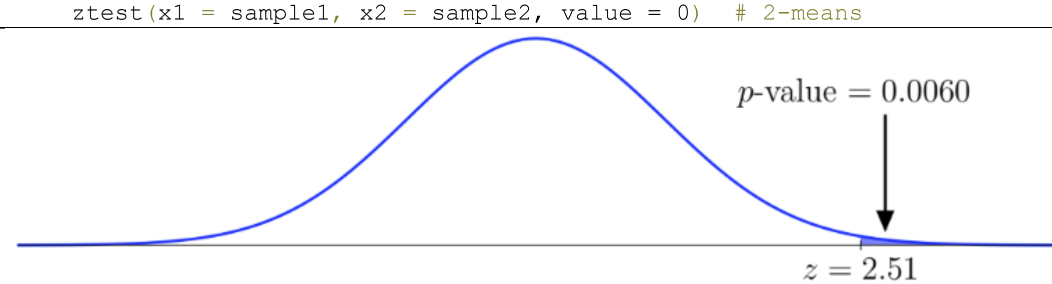
z-score = -0.5802
p-value = 0.5618 (two-tailed)

6) Calculate p-value

p-value

TI-84: DISTR 2: normalcdf(z_data, 99999999) = p
Python:

```
ztest(x1 = scores['Exam1'], H0_value = 86)     # 1-mean
ztest(x1 = sample1, x2 = sample2, value = 0)   # 2-means
```



7) Conclusion	
Statistical Decision	Reject the null hypothesis (supporting the alternative hypothesis) using a test below.
Conclusions of p-test	If p–value < α ⇒ Reject H_0 in favor of H_a. If p–value ≥ α ⇒ Fail to Reject H_0.
Conclusions of mean test	If significance level (α) is less than 5% ⇒ Reject H_0 in favor of H_a. If confidence level (c) is greater than 95% ⇒ Reject H_0 in favor of H_a. If test statistic is greater than (right-tailed) the critical value, $z_{data} > z^*$ ⇒ Reject H_0.
Conclusions of Confidence Interval for μ / z interval	Reject the null hypothesis if the test statistic falls in the **rejection region** otherwise, fail to reject the null hypothesis. If confidence interval is between 5% and 95%, meaning ($\pm$5%) ⇒ Reject H_0.

Hypothesis Testing Error Types

Ideally, a statistical test should have a low significance level (α) and high power (1−β).

Type I Error (α): False Positive
Type II Error (β): False Negative

		H_0 true	H_0 false
Decision	Reject H_0	Type I error Prob $= \alpha$	Correct decision
	Fail to reject H_0	Correct decision	Type II error Prob $= \beta$