STATISTICS FOR MANAGERS

Practice Book

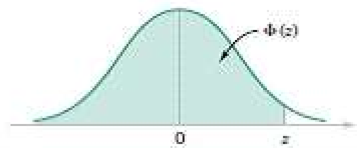
OCTOBER 7, 2024

APTECH SOLUTIONS LLC
Ohio

Contents

Z-table

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du$$



z	−0.09	−0.08	−0.07	−0.06	−0.05	−0.04	−0.03	−0.02	−0.01	−0.00	z
−3.9	0.000033	0.000034	0.000036	0.000037	0.000039	0.000041	0.000042	0.000044	0.000046	0.000048	−3.9
−3.8	0.000050	0.000052	0.000054	0.000057	0.000059	0.000062	0.000064	0.000067	0.000069	0.000072	−3.8
−3.7	0.000075	0.000078	0.000082	0.000085	0.000088	0.000092	0.000096	0.000100	0.000104	0.000108	−3.7
−3.6	0.000112	0.000117	0.000121	0.000126	0.000131	0.000136	0.000142	0.000147	0.000153	0.000159	−3.6
−3.5	0.000165	0.000172	0.000179	0.000185	0.000193	0.000200	0.000208	0.000216	0.000224	0.000233	−3.5
−3.4	0.000242	0.000251	0.000260	0.000270	0.000280	0.000291	0.000302	0.000313	0.000325	0.000337	−3.4
−3.3	0.000350	0.000362	0.000376	0.000390	0.000404	0.000419	0.000434	0.000450	0.000467	0.000483	−3.3
−3.2	0.000501	0.000519	0.000538	0.000557	0.000577	0.000598	0.000619	0.000641	0.000664	0.000687	−3.2
−3.1	0.000711	0.000736	0.000762	0.000789	0.000816	0.000845	0.000874	0.000904	0.000935	0.000968	−3.1
−3.0	0.001001	0.001035	0.001070	0.001107	0.001144	0.001183	0.001223	0.001264	0.001306	0.001350	−3.0
−2.9	0.001395	0.001441	0.001489	0.001538	0.001589	0.001641	0.001695	0.001750	0.001807	0.001866	−2.9
−2.8	0.001926	0.001988	0.002052	0.002118	0.002186	0.002256	0.002327	0.002401	0.002477	0.002555	−2.8
−2.7	0.002635	0.002718	0.002803	0.002890	0.002980	0.003072	0.003167	0.003264	0.003364	0.003467	−2.7
−2.6	0.003573	0.003681	0.003793	0.003907	0.004025	0.004145	0.004269	0.004396	0.004527	0.004661	−2.6
−2.5	0.004799	0.004940	0.005085	0.005234	0.005386	0.005543	0.005703	0.005868	0.006037	0.006210	−2.5
−2.4	0.006387	0.006569	0.006756	0.006947	0.007143	0.007344	0.007549	0.007760	0.007976	0.008198	−2.4
−2.3	0.008424	0.008656	0.008894	0.009137	0.009387	0.009642	0.009903	0.010170	0.010444	0.010724	−2.3
−2.2	0.011011	0.011304	0.011604	0.011911	0.012224	0.012545	0.012874	0.013209	0.013553	0.013903	−2.2
−2.1	0.014262	0.014629	0.015003	0.015386	0.015778	0.016177	0.016586	0.017003	0.017429	0.017864	−2.1
−2.0	0.018309	0.018763	0.019226	0.019699	0.020182	0.020675	0.021178	0.021692	0.022216	0.022750	−2.0
−1.9	0.023295	0.023852	0.024419	0.024998	0.025588	0.026190	0.026803	0.027429	0.028067	0.028717	−1.9
−1.8	0.029379	0.030054	0.030742	0.031443	0.032157	0.032884	0.033625	0.034379	0.035148	0.035930	−1.8
−1.7	0.036727	0.037538	0.038364	0.039204	0.040059	0.040929	0.041815	0.042716	0.043633	0.044565	−1.7
−1.6	0.045514	0.046479	0.047460	0.048457	0.049471	0.050503	0.051551	0.052616	0.053699	0.054799	−1.6
−1.5	0.055917	0.057053	0.058208	0.059380	0.060571	0.061780	0.063008	0.064256	0.065522	0.066807	−1.5
−1.4	0.068112	0.069437	0.070781	0.072145	0.073529	0.074934	0.076359	0.077804	0.079270	0.080757	−1.4
−1.3	0.082264	0.083793	0.085343	0.086915	0.088508	0.090123	0.091759	0.093418	0.095098	0.096801	−1.3
−1.2	0.098525	0.100273	0.102042	0.103835	0.105650	0.107488	0.109349	0.111233	0.113140	0.115070	−1.2

Quiz-2

Problem-1:

Mike has been delayed in going to the annual sales event at one of his favorite apparel stores. His friend has just texted him telling him that there are only 20 shirts left, of which 8 are size M, 10 are size L, and 2 are size XL. Also, 3 shirts are white, 5 are blue, and the remaining are mixed colors. Mike is interested in getting a white or blue shirt in size L.

Let:

A = Getting a white or blue shirt

B = Getting a shirt of size L

What is P(A)? Report your answer as a decimal up to 2 decimal places.

Answers:

Numbers from the problem

- Total shirts = 20
- White shirts = 3
- Blue shirts = 5
- Size L shirts = 10

Finding the Probabilities:

1. **P(A):** Probability of getting a white or blue shirt
    - Since white and blue shirts are mutually exclusive (a shirt can't be both white and blue at the same time), we can add their individual probabilities.
    - P(A) = P(White) + P(Blue) = (3/20) + (5/20) = 8/20 = 0.4

Problem-2

A patient visits her doctor with concerns about her blood pressure. If the systolic blood pressure exceeds 150, the patient is considered to have high blood pressure and medication may be prescribed. The problem is that there is considerable variation in the patient's systolic BP readings during the day.

If a patient's SBP readings during a given day have a normal distribution with a mean of 160 mmHg and a standard deviation 20 mmHg, what is the probability that a single measurement will fail to detect that the patient has high blood pressure?

Report your answer as a decimal to 4 places.

Numbers from the problem

- Mean (μ) = 160 mmHg
- Standard Deviation (σ) = 20 mmHg

- Threshold (X) = 150 mmHg

Using the Z-Score Formula

The Z-score is a standardized value that tells us how many standard deviations a data point is from the mean. It's calculated using the formula:

$Z = (X - \mu) / \sigma$

Calculating the Z-score:

- $Z = (150 - 160) / 20 = -0.5$

Excel Formula:

- =NORM.S.DIST(-0.5, TRUE)

Result:

- The result of the formula will be approximately 0.3085.

Therefore, the probability that a single measurement will fail to detect that the patient has high blood pressure is 0.3085.

Problem-3:

A manager of a popular retail store knows that the distribution of purchase amounts by its customers is approximately normal with a mean of $30 and a standard deviation of $9. Below you will find normal probability and percentile calculations related to the customer purchase amounts.

What is the probability that a randomly selected customer will spend between $20 and $35? Report your answer as a decimal with no more than 2 decimal places.

Solution:

Looks like a normal distribution problem and to find the probability of a value falling within a specific range.

Number in problem:

- Mean (μ) = $30
- Standard Deviation (σ) = $9
- Lower bound (X1) = $20
- Upper bound (X2) = $35

Using the Z-Score Formula

$Z = (X - \mu) / \sigma$

Calculating Z-scores:

- $Z1 = (20 - 30) / 9 \approx -1.11$

- $Z2 = (35 - 30) / 9 \approx 0.56$

Excel Formulas:

- =NORM.S.DIST(0.56, TRUE) - NORM.S.DIST(-1.11, TRUE)

This formula calculates the cumulative probability for Z-scores less than 0.56 and subtracts the cumulative probability for Z-scores less than -1.11, giving us the probability between the two Z-scores.

Result:

- The result of the formula will be approximately 0.41.

Therefore, the probability that a randomly selected customer will spend between $20 and $35 is 0.41.

Problem-4:

Wendy's fast-food restaurant sells hamburgers and chicken sandwiches. Suppose that on a typical weekday, the demand for hamburgers is normally distributed with a mean of 450 and standard deviation of 80 and the demand for chicken sandwiches is normally distributed with a mean of 120 and standard deviation of 30. Use this information to answer the following questions.

How many chicken sandwiches must the restaurant stock so that there is only a 5% chance of running out of sandwiches? Report your answer with no decimals.

Solution:

Numbers from problem:

- **Mean demand for chicken sandwiches (μ)** = 120
- **Standard deviation (σ)** = 30
- **Probability of running out (P)** = 5% (0.05)

Z-score formula for 95%:

=NORM.S.INV(0.95)

- X is the number of chicken sandwiches to stock.
- μ\muμ is the mean demand (120).
- ZZZ is the Z-score (1.645).
- σ\sigmaσ is the standard deviation (30).

Forumula - $X=\mu+Z\cdot\sigma$

$X=120+(1.645\times30)=120+49.35=169.35$ rounding to 170

Quiz::3

Question 12:

Suppose that the average weekly earnings for employees in general automotive repair shops is \$450, and that the standard deviation for the weekly earnings for such employees is \$50. A sample of 100 such employees is selected at random. Find the probability that the mean of the sample is between \$445 and \$455. Report this probability as a decimal with two places.

Solution:

SRS problem. Central limit theorem as sample size is large enough ($n >= 30$)

Numbers in problem

- Population mean (μ) = \$450
- Population standard deviation (σ) = \$50
- Sample size (n) = 100
- Lower bound (X1) = \$445
- Upper bound (X2) = \$455

Standard Error formula: $SE = \sigma / \sqrt{n}$

$SE = 50 / \sqrt{100} = 5$

For Z score $Z = (X - \mu) / SE$

$Z1 = (445-450)/5 = -1$

$Z2 = (450-445)/5 = 1$

Probability of z-scre =NORM.S.DIST(1, TRUE) - NORM.S.DIST(-1, TRUE) = 0.68 (ROUNDED)

Therefore, the probability that the mean of the sample is between \$445 and \$455 is 0.68.

Question 13:

Recently some complaints have been filed with the consumer bureau against a certain pizzeria for allegedly failing to meet their advertising claim that each of their large pepperoni pizza is topped with 2 ounces of pepperoni on average. A city inspector in charge of the case decides to buy 25 large pepperoni pizzas and weigh each one. Assume that the weight of the pepperoni varies normally with a standard deviation of 0.05 oz. What is the probability that the average weight of pepperoni of 1.98 oz is seen?

Round your answer to three decimal places.

Solution:

Numbers in problem:

- Population mean (μ) = 2 oz
- Population standard deviation (σ) = 0.05 oz

- Sample size (n) = 25
- Sample mean (X1) = 1.98 oz

Standard Error formula: SE = $\sigma / \sqrt{n}$

$SE = 0.05 / \sqrt{25} = 0.01$

For Z score $Z = (X - \mu) / SE$

$Z = (1.98-2)/0.01 = -2$

Probability of z-scre =NORM.S.DIST(1, TRUE) - NORM.S.DIST(-2, TRUE) = 0.0228 = 0.023(ROUNDED)

Therefore, the probability is 0.023

Question 15:

A car manufacturer conducted a study by randomly sampling and interviewing 1,000 consumers in a new target market. The goal of the study was to determine if consumers would consider purchasing this brand of car.

Management has already determined that the company will enter this segment. In the sample, 93 consumers exhibited what the company considered strong brand liking.

What is the probability of observing 93 consumers (or fewer) in 1000 if the true proportion is 10%?

Report the probability as a decimal rounded to two decimal places.

Solution:

Solution:

Numbers in problem:

- Probabilty of success(p) =0.10
- Number of scucess (x) = 93
- Sample size (n) = 1000

Approximating with Normal Distribution

We can use the normal approximation to the binomial distribution since np = 1000 * 0.10 = 100 and n(1-p) = 1000 * 0.90 = 900, both of which are greater than 5.

Mean and Standard Deviation:

- Mean (μ) = np = 100
- Standard Deviation (σ) = $\sqrt{np(1-p)} = \sqrt{100 * 0.90} \approx 9.49$

Standardizing the Value

We need to standardize the number of successes using the Z-score formula:

- $Z = (x - \mu) / \sigma = (93 - 100) / 9.49 \approx -0.74$

Finding the Probability Excel Formula:

- =NORM.S.DIST(-0.74, TRUE)

This formula calculates the cumulative probability for a standard normal distribution up to the specified Z-score. The TRUE argument indicates that we want the cumulative probability.

Result**:**

- The result of the formula will be approximately 0.23.

Therefore, the probability of observing 93 consumers (or fewer) in 1000 if the true proportion is 10% is 0.23.

Quiz-5

Question-1

The analyst gets to choose the significance level . It is typically chosen to be 0.50, but it is occasionally chosen to be 0.01.

False

- **Explanation:** The significance level (α) is typically chosen to be a small value, such as 0.05 or 0.01. This reflects the willingness to accept a certain level of Type I error (rejecting a true null hypothesis). Choosing a significance level of 0.50 would be very high, indicating a high tolerance for Type I errors.

Question-2

Sample evidence is statistically significant at the level only if the p–value is larger than .

False

- **Explanation:** Sample evidence is statistically significant at the α level only if the p-value is *smaller* than α. A p-value larger than α indicates that the observed data is likely to occur under the null hypothesis.

Question-3

A professor of statistics doubts the claim that the proportion of voters in a particular party in Michigan is at most 45%. To test the claim, the hypotheses:, , should be used.

True

- **Explanation:** The null hypothesis should represent the "no effect" or "status quo" scenario. In this case, the claim is that the proportion is *at most* 45%. Therefore, the null hypothesis should be H_0: $p \leq 0.45$, and the alternative hypothesis would be H_1: $p > 0.45$.

Question-4

A null hypothesis is a statement about the value of a population parameter. It is usually the current thinking, or "status quo".

True

- **Explanation:** The null hypothesis is a statement about a population parameter. It is typically a statement of "no difference" or "no effect."

Question- 5

The alternative hypothesis, or research hypothesis, is the hypothesis that the analyst is attempting to prove.

True

- **Explanation:** The alternative hypothesis is the research hypothesis, representing the claim that the analyst is trying to prove.

Question-6

A one-tailed alternative is one that is supported by evidence in either direction.

False

- **Explanation:** A one-tailed alternative hypothesis is one that specifies a direction of the effect. For example, H_1: $p > 0.5$ or H_1: $p < 0.5$. A two-tailed alternative hypothesis does not specify a direction, such as H_1: $p \neq 0.5$.

Question-7

The p-value of a test is the probability of observing a test statistic at least as extreme as the one computed given that the null hypothesis is true.

True

- **Explanation:** The p-value is the probability of observing a test statistic as extreme or more extreme than the one calculated, assuming the null hypothesis is true

Question-8

Based on the 2000 Census, the proportion of the California population aged 15 years old or older who are married is $p = 0.524$. Suppose $n = 1000$ persons are to be sampled from this population and the sample proportion of married persons is to be calculated. In our sample, 500 out of 1000 were married.

Conduct a hypothesis test whether or not the Census data appears to be true now. State your null and alternative hypothesis, what test you are using, assumptions, p-value, decision and conclusions.

Solution

- **Random Sample:** The sample of 1000 individuals was drawn randomly from the California population.
- **Large Sample Size:** Both np and n(1-p) are greater than 10. In this case, np = 1000 * 0.524 = 524 and n(1-p) = 1000 * 0.476 = 476, both of which are greater than 10.
- **Null Hypothesis (H_0):** p = 0.524 (The proportion of married persons in California remains unchanged from the 2000 Census)
- **Alternative Hypothesis (H_1):** $p \neq 0.524$ (The proportion of married persons in California has changed since the 2000 Census)

Test: Using z-test for proportions

- Population proportion (p) = 0.524
- Sample size (n) = 1000
- Sample proportion (phat) = 500/1000 = 0.5

z = (phat - p) / Sqrt(p(1-p)/n)

z = (0.5-0.524)/Sqrt(0.524(1-0.524)/1000)
= -1.39

So p-value from table 0.164

The significance level (α) is typically chosen to be a small value, such as 0.05 or 0.01 but we see the p is 0.164 > 0.05. **We fail to reject the null hypothesis**.

There is not enough evidence to conclude that the proportion of married persons in California has changed significantly from the 2000 Census data.

Based on the sample, the observed proportion of married persons is not statistically different from the population proportion reported in the 2000 Census.

Question-9

According to the CDC, the mean height of adults ages 20 and older is about 66.5 inches (69.3 inches for males, 63.8 inches for females). Let's test if the mean height of our sample data, n = 60, is significantly different than 66.5 inches. The observed average height in our sample is 68 inches and the standard deviation is 5.5 inches.

State your null and alternative hypothesis, what test you are using, assumptions, p-value, decision and conclusions.

Solution

- Random Sample: The sample of 60 individuals was drawn randomly from the population of adults ages 20 and older.
- Normality: The population of heights is normally distributed. If the sample size is large (n >= 30), the Central Limit Theorem suggests that the sampling distribution of the mean will be approximately normal, even if the population is not normally distributed.
- Null Hypothesis (H0): mu = 66.5 inches (The mean height of the sample is not significantly different from the population mean)
- Alternative Hypothesis (H1): mu ≠ 66.5 inches (The mean height of the sample is significantly different from the population mean.

Since we are testing a population mean and the population standard deviation is unknown. We will use a t-test

t = (xbar - mu)/(s/sqrt(n))

t = (68 - 66.5) / (5.5 / sqrt(60))

t = 2.07

df = n-1 = 601 - 1 = 59

p-value is 0.0403 < 0.05 (typical significance value α). So, If we choose a significance level (α) of 0.05, since the p-value (0.043) is less than α, we reject the null hypothesis.

There is enough evidence to conclude that the mean height of the sample is significantly different from the population mean of 66.5 inches. The observed sample mean of 68 inches suggests that the individuals in this sample are taller than the average adult population.

Mean/Proportions

Keywords Mapping

Keyword or Phrase	Indicates	Explanation
Mean, average, median, mode	Averages	These terms typically refer to numerical data, such as measurements, scores, or values.
Proportion, percentage, rate	Proportions	These terms typically refer to categorical data, such as yes/no, success/failure, or categories.
Per, out of, ratio	Proportions	These terms often indicate a comparison between two quantities, which can be expressed as a proportion or percentage.
Sum, total, aggregate	Averages	These terms suggest that the data is being added together to find a central value.
Success, failure, yes/no, true/false	Proportions	These terms indicate categorical data that can be counted as successes or failures.
Continuous data (e.g., weight, height, time)	Averages	Continuous data can be measured on a continuous scale and is typically analyzed using averages.
Categorical data (e.g., color, gender, opinion)	Proportions	Categorical data is divided into categories and is typically analyzed using proportions or percentages.

For Practice

1. Suppose the Acme Drug Company develops a new drug, designed to prevent colds. The company states that the drug is equally effective for men and women. To test this claim, they choose a a simple random sample of 100 women and 200 men from a population of 100,000 volunteers. At the end of the study, 38% of the women caught a cold; and 51% of the men caught a cold. Which case is this?
   Two sample proportion

2. Recently some complaints have been filed with the consumer bureau against a certain pizzeria for allegedly failing to meet their advertising claim that each of their large pepperoni pizza is topped with 2 ounces of pepperoni on average. A city inspector in charge of the case decides to buy 25 large pepperoni pizzas and weigh each one. The weight of the pepperoni varies normally. Our sample of 25 yields a mean of 1.5 oz with standard deviation of 0.05 oz. Which case is this?
   One sample mean

3. Based on the 2000 Census, the proportion of the California population aged 15 years old or older who are married is $p = 0.524$. Suppose $n = 1000$ persons are to be sampled from this population and the sample proportion of married persons is to be calculated. In our sample 500 out of 1000 were married. Which case is this?
   One sample proportion

4. A concern was raised in Australia that the percentage of deaths of Aboriginal prisoners was higher than the percent of deaths of non-Aboriginal prisoners, which is 0.27%. A sample of six years (1990-1995) of data was collected, and it was found that out of 14,495 Aboriginal prisoners, 51 died ("Indigenous deaths in," 1996). Which case is this?
   One sample proportion

5. Vital capacity is a measure of the amount of air that someone can exhale after taking a deep breath. One might expect that musicians who play brass instruments would have greater vital capacities, on average, then would other persons of the same age, sex, and height. Vital capacity follows a normal distribution. The average of the vital capacity for seven brass players is 4.83 with standard deviation of 0.435 and for a sample of 5 other people the average is 4.74 with standard deviation of 0.351. Which case is this?
   Two independent samples – difference in averages

6. Under certain conditions, electrical stimulation of a beef carcass will improve the tenderness of the meat. In one study of this effect, beef carcasses were split in half, one side was subjected to a brief electrical current and the other side is an untreated control. For each side, a steak was cut and tested in various ways for tenderness. In one test, the experimenter obtained a specimen of connective tissue (collagen) from the steak and determined the temperature at which the tissue would shrink: a tender piece of meat tends to yield a low collagen shrinkage temperature. Which case is this?
   Paired Samples – average problem

7. In an election, two different political parties are running. The number of people who voted for each party is measured. We are interested in whether the proportions are different (i.e. one party might have more support than the other). For political party one, we have 500 voters in our area, 300 of them voted. For political party two, we had 600 voters in our area and 350 of them voted. Which case is this?
   Two sample proportion

8. A hospital has a random sample of cholesterol measurements for 100 men the mean cholesterol was 220 with a standard deviation of 10. These patients were seen for issues other than cholesterol. They were not taking any medications for high cholesterol. The hospital wants to know if the unknown mean cholesterol for patients is different from a goal level of 200 mg. Which case is this?
   One sample mean

9. A cross-over trial investigated whether eating oat bran lowered serum cholesterol levels. Fourteen (14) individuals were randomly assigned a diet that included either oat bran or corn flakes. After two weeks on the initial diet, serum cholesterol were measured and the participants were then "crossed-over" to the alternate diet. After two-weeks on the second diet, cholesterol levels were once again recorded. Data appear below. The variable CORNFLK in the table represents cholesterol level (mmol/L) of the participant on the corn flake diet. The variable OATBRAN represents the participant's cholesterol on the oat bran diet. We are interested in whether eating oat bran results in lower cholesterol compared to the control (corn flakes). You are given the mean and standard deviation of each group as well as the mean and standard deviation of the difference. Which case is this?
   Paired Samples – average problem

10. We have students who speak English as their first language and students who do not. All students take a reading test. Our two groups are the native English speakers and the non-native speakers. Our measurements are the test scores. Our idea is that the mean test scores for the underlying populations of native (n=124, mean = 85, standard deviation of 8) and non-native (n=149, mean = 75, standard deviation of 12) English speakers are not the same. We want to know if the mean score for the population of native English speakers is different from the people who learned English as a second language. Which case is this?
    Two independent samples – difference in averages

Hypothesis Tests

Tabular view

Hypothesis Test	Statistical Formula	Excel Formulas	Steps
Hypothesis Test for a Population Mean	t-test (if σ is unknown) or z-test (if σ is known)	T.TEST(array1, array2, tails, type) or =Z.TEST(array, μ, σ)	1. State hypotheses (H0 and Ha) 2. Calculate test statistic 3. Determine critical value or p-value 4. Make a decision
Hypothesis Test for a Population Mean, Broken Down by a Categorical Variable	ANOVA	F.TEST(array1, array2)	1. State hypotheses (H0 and Ha) 2. Calculate F-statistic 3. Determine critical value or p-value 4. Make a decision
Hypothesis Test for a Population Total	t-test or z-test (depending on σ)	Same as for population mean	1. State hypotheses (H0 and Ha) 2. Calculate test statistic 3. Determine critical value or p-value 4. Make a decision
Hypothesis Test for a Population Standard Deviation	Chi-square test	CHISQ.TEST(array1, array2)	1. State hypotheses (H0 and Ha) 2. Calculate chi-square statistic 3. Determine critical value or p-value 4. Make a decision
Hypothesis Test for a Population Proportion Having Some Property of Interest	z-test	Z.TEST(array, μ, σ)	1. State hypotheses (H0 and Ha) 2. Calculate test statistic 3. Determine critical value or p-value 4. Make a decision
Hypothesis Test for the Difference Between Two Means from Independent Samples	t-test (if σ1 ≠ σ2) or pooled t-test (if σ1 = σ2)	T.TEST(array1, array2, tails, type) (for pooled t-test, use =T.TEST(array1, array2, tails, 2))	1. State hypotheses (H0 and Ha) 2. Calculate test statistic 3. Determine critical value or p-value 4. Make a decision
Hypothesis Test for the Difference Between Two	Paired t-test	T.TEST(array1, array2, tails, 1)	1. State hypotheses (H0 and Ha) 2. Calculate test

Means When the Two Samples Are Paired in Some Natural Way			statistic 3. Determine critical value or p-value 4. Make a decision
Hypothesis Test for the Difference Between Two Population Proportions	z-test	Z.TEST(array1, p, σ) (where p is the pooled proportion)	1. State hypotheses (H0 and Ha) 2. Calculate test statistic 3. Determine critical value or p-value 4. Make a decision
Hypothesis Test for the Ratio of Two Variances from Two Independent Samples	F-test	F.TEST(array1, array2)	1. State hypotheses (H0 and Ha) 2. Calculate F-statistic 3. Determine critical value or p-value 4. Make a decision

- **Tails:** For a one-tailed test, use 1 or 2. For a two-tailed test, use 2.
- **Type:** For a paired t-test, type = 1. For a two-sample t-test with equal variances, type = 2. For a two-sample t-test with unequal variances, type = 3.
- **Array1 and array2:** These are the ranges of data for the two samples.
- **μ:** The hypothesized population mean.
- **σ:** The population standard deviation.
- **p:** The hypothesized population proportion.

Practice Problems

Hypothesis Test for a Population Mean

Problem 1: A teacher claims that the average IQ of students in her class is higher than the national average of 100. To test this claim, she administers an IQ test to her class of 25 students. The sample mean IQ is 105 with a standard deviation of 10. Is there sufficient evidence to support the teacher's claim at a 0.05 significance level?

- **Hypothesis:**
  - H0: $\mu = 100$
  - Ha: $\mu > 100$ (one-tailed test)
- **Reason for Selection:** We are testing a claim about the population mean based on a single sample.

Hypothesis Test for a Population Mean, Broken Down by a Categorical Variable (ANOVA)

Problem 2: A researcher wants to compare the average salaries of three different professions: doctors, lawyers, and engineers. She collects salary data for a random sample of 20 doctors, 25 lawyers, and 15 engineers. Is there a significant difference in the average salaries among these professions?

- **Hypothesis:**
  - H0: $\mu 1 = \mu 2 = \mu 3$
  - Ha: At least one mean is different
- **Reason for Selection:** We are comparing the means of three or more groups.

Hypothesis Test for a Population Total

Problem 3: A company wants to estimate the total number of defective products produced in a month. They randomly sample 100 products and find that 5 of them are defective. Based on this sample, can we estimate the total number of defective products in the month with a 95% confidence interval?

- **Hypothesis:**
  - This is not a hypothesis test but rather a confidence interval estimation problem.
- **Reason for Selection:** The goal is to estimate a population total based on a sample.

Hypothesis Test for a Population Standard Deviation

Problem 4: A quality control engineer wants to test if the standard deviation of the weight of a certain product has changed from its historical value of 0.5 grams. She measures the weight of 20 randomly selected products and finds a sample standard deviation of 0.6 grams. Is there sufficient evidence to conclude that the standard deviation has changed at a 0.01 significance level?

- **Hypothesis:**
  - H0: $\sigma = 0.5$

- Ha: $\sigma \neq 0.5$ (two-tailed test)

- **Reason for Selection:** We are testing a claim about the population standard deviation.

Hypothesis Test for a Population Proportion Having Some Property of Interest

Problem 5: A political pollster wants to estimate the proportion of voters who support a new tax reform. They survey 1000 randomly selected voters and find that 600 of them support the tax reform. Is there sufficient evidence to conclude that the majority of voters support the tax reform at a 0.05 significance level?

- **Hypothesis:**
  - H0: $p = 0.5$
  - Ha: $p > 0.5$ (one-tailed test)
- **Reason for Selection:** We are testing a claim about a population proportion.

Hypothesis Test for the Difference Between Two Means from Independent Samples

Problem 6: A researcher wants to compare the average test scores of students who are taught using a new teaching method to those who are taught using the traditional method. She randomly assigns 30 students to each group and conducts the experiment. Is there a significant difference in the average test scores between the two groups?

- **Hypothesis:**
  - H0: $\mu 1 - \mu 2 = 0$
  - Ha: $\mu 1 - \mu 2 \neq 0$ (two-tailed test)
- **Reason for Selection:** We are comparing the means of two independent groups.

Hypothesis Test for the Difference Between Two Means When the Two Samples Are Paired in Some Natural Way

Problem 7: A company wants to test the effectiveness of a new weight loss program. They measure the weight of 20 participants before and after the program. Is there a significant difference in the average weight loss?

- **Hypothesis:**
  - H0: $\mu d = 0$
  - Ha: $\mu d > 0$ (one-tailed test)
- **Reason for Selection:** The data is paired (before and after measurements).

Hypothesis Test for the Difference Between Two Population Proportions

Problem 8: A political pollster wants to compare the proportion of men and women who support a new political candidate. They survey 500 men and 500 women and find that 300 men and 250 women support the candidate. Is there a significant difference in the proportion of men and women who support the candidate?

- **Hypothesis:**
    - H0: $p1 - p2 = 0$
    - Ha: $p1 - p2 \neq 0$ (two-tailed test)
- **Reason for Selection:** We are comparing two proportions.

Hypothesis Test for the Ratio of Two Variances from Two Independent Samples

Problem 9: A manufacturing company wants to compare the variability of two production lines. They measure the weight of 20 products from each line. Is there a significant difference in the variance of the weights between the two lines?

- **Hypothesis:**
    - H0: $\sigma 1^2 / \sigma 2^2 = 1$
    - Ha: $\sigma 1^2 / \sigma 2^2 \neq 1$ (two-tailed test)
- **Reason for Selection:** We are comparing the variances of two independent groups.

Compiled for practice 😊 Sumanth Krishna, Ambati