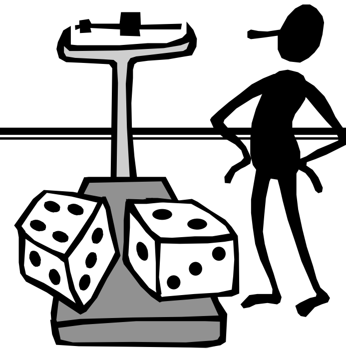
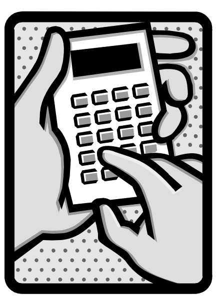
Chapter 1: Exploring Data



Key Vocabulary:

- individuals
- variables
- categorical variable
- quantitative variable
- distribution
- range
- spread
- frequency
- outlier
- center
- shape
- skewed left
- skewed right
- symmetric
- dot plot
- histogram
- stemplot
- split stems
- back-to-back stemplot
- time plot
- mean
- Σ
- $\overline{x}$
- nonresistant
- median
- resistant
- quartiles
- Q_1, Q_3
- IQR
- five-number summary
- minimum
- maximum
- boxplot
- modified boxplot
- standard deviation
- variance

Calculator Skills:



- round( )
- sortA( )
- sum
- mean
- 1-Var Stats
- Float
- ENTRY
- ANS
- ZoomStat
- TRACE
- WINDOW

1.1 Displaying Distributions with Graphs (pp.4-30)

1. In statistics, what is meant by *individuals*?

2. In statistics, what is meant by a *variable*?

3. What is meant by *exploratory data analysis*?

4. What is the difference between a *categorical variable* and a *quantitative variable*?

5. When is it useful to use a bar chart?

6. When is it useful to use a pie chart?

7. What is meant by a *distribution*?

8. Define *range*:

9. When is it better to use a *histogram* rather than a *dotplot*?

10. What is meant by *frequency* in a histogram?

11. When setting a window for constructing a histogram on the TI-83:
    a. What is the significance of Xscl?
    b. How do you choose the values of Xmin and Xmax?
    c. What is the significance of Ymax?

12. Define *outlier*.

13. If a distribution is *symmetric*, what does its histogram look like?

14. If a distribution is *skewed right*, what does its histogram look like?

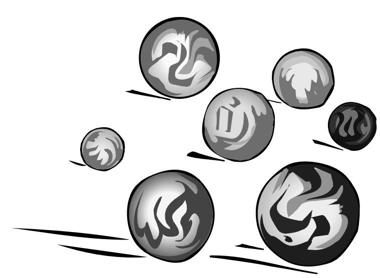
15. If a distribution is *skewed left*, what does its histogram look like?

16. How is the *stemplot* of a distribution related to its histogram?

17. When is it advantageous to split stems on a stemplot?

18. What is the purpose of a *back-to-back stemplot*?

19. When is it useful to construct a *time plot*?



1.2 Describing Distributions with Numbers (pp.30-51)

1. In statistics, what is the most common measurement of center?

2. Explain how to calculate the *mean*, $\bar{x}$.

3. Explain how to calculate the *median*, *M*.

4. Explain why the median is *resistant* to extreme observations, but the median is *nonresistant*.

5. In statistics, what is meant by *spread*?

6. Explain how to calculate Q_1 and Q_3.

7. What is the *five-number summary*?

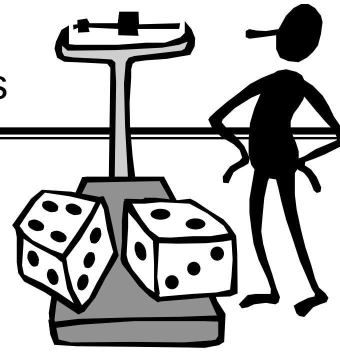
8. What does *standard deviation* measure?

9. What is the relationship between *variance* and *standard deviation*?

10. When does *standard deviation* equal zero?

11. Is *standard deviation* resistant or nonresistant to extreme observations? Explain.

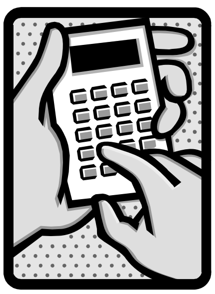
Chapter 2: The Normal Distributions



Key Vocabulary:

- density curve
- μ mu
- σ sigma
- outcomes
- simulation
- normal curve
- normal distribution
- inflection point
- 68-95-99.7 rule
- percentile
- $N(\mu, \sigma)$
- standardized value
- z-scores
- standard normal distribution
- normal probability plot

Calculator Skills:



- randInt
- $X[35, 185]_{25}$
- $Y[-.01, .02]_{.01}$
- rand
- ShadeNorm(lowerbound, upperbound, μ, σ)
- normalpdf(x, μ, σ)
- normalcdf(lowerbound, upperbound, μ, σ)
- EE (1E99 and -1E99)
- invNorm(area, μ, σ)

2.1 Density Curves and the Normal Distributions (pp.64-82)

1. What is a *density curve*?

2. What does the area under a *density curve* represent?

3. Where is the median of a *density curve* located?

4. Where is the mean of a *density curve* located?

5. What is a *uniform distribution*?

6. What is the difference between the *randInt* and *rand* commands on the TI-83?

7. How would you describe the shape of a *normal curve*? Draw several examples.

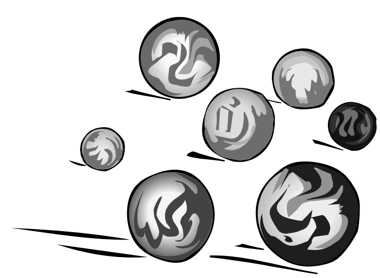
8. Where on the *normal curve* are *inflection points* located?

9. Explain the *68-95-99.7 Rule*.

10. What is a *percentile*?

11. Is there a difference between the 80^{th} percentile and the top 80%? Explain.

12. Is there a difference between the 80^{th} percentile and the lower 80%? Explain.



2.2 Standard Normal Calculations (pp.83-100)

1. Explain how to *standardize* a variable.

2. What is the purpose of standardizing a variable?

3. What is the *standard normal distribution*?

4. What information does the *standard normal table* give?

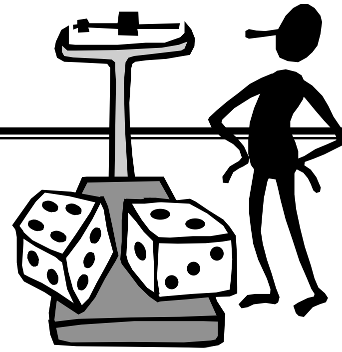
5. How do you use the standard normal table (Table A) to find the area under the standard normal curve to the left of a given *z-value*? Draw a sketch.

6. How do you use Table A to find the area under the standard normal curve to the right of a given *z-value*? Draw a sketch.

7. How do you use Table A to find the area under the standard normal curve between two given *z-values*? Draw a sketch.

8. Describe two methods for assessing whether or not a distribution is *approximately normal*.

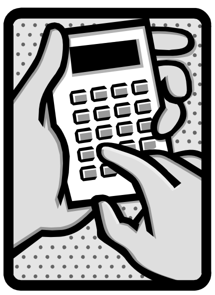
Chapter 3: Examining Relationships



Key Vocabulary:

- response variable
- explanatory variable
- independent variable
- dependent variable
- scatterplot
- positive association
- negative association
- linear
- correlation
- r-value
- regression line
- mathematical model
- least-squares regression line
- $\hat{y}$ "y-hat"
- SSM
- SSE
- r^2
- coefficient of determination
- residuals
- residual plot
- influential observation

Calculator Skills:



- seq(X,X,min,max,scl)
- 2-Var Stats
- sum
- Diagnostic On
- $\bar{x}, s_x, \bar{y}, s_y$
- Clear All Lists
- residual plot

3.1 Scatterplots (pp.107-127)

1. What is the difference between a *response variable* and an *explanatory variable*?

2. How are response and explanatory variables related to *dependent* and *independent* variables?

3. When is it appropriate to use a *scatterplot* to display data?

4. Which variable always appears on the horizontal axis of a scatterplot?

5. Explain the difference between a *positive association* and a *negative association*.

3.2 Correlation (pp.128-136)

1. What does *correlation* measure?

2. Explain why two variables must both be *quantitative* in order to find the *correlation* between them.

3. What is true about the relationship between two variables if the *r-value* is:
   a. Near 0?
   b. Near 1?
   c. Near -1?
   d. Exactly 1?
   e. Exactly -1?

4. Is *correlation* resistant to extreme observations? Explain.

5. What does it mean if two variables have *high correlation*?

6. What does it mean if two variables have *weak correlation*?

7. What does it mean if two variables have *no correlation*?

3.3 Least-Squares Regression (pp.137-163)

1. In what way is a *regression line* a *mathematical model*?

2. What is a *least-squares regression line*?

3. What is the formula for the equation of the *least-squares regression line*?

4. How is *correlation* related to *least-squares regression*?

5. What is the formula for calculating the *coefficient of determination*?

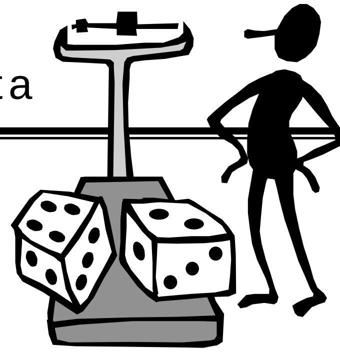
6. The r^2 value shows how much of the variation in one variable can be accounted for by the linear relationship with the other variable. If $r^2 = 0.95$, what can be concluded about the relationship between x and y?

7. Define *residual*.

8. If a *least-squares regression line* fits the data well, what characteristics should the *residual plot* exhibit?

9. What is meant by an *influential observation*?

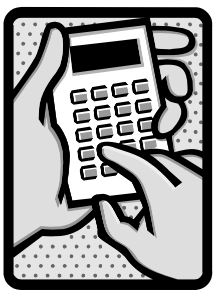
Chapter 4: More on Two-Variable Data



Key Vocabulary:

- exponential function
- power function
- linear growth
- exponential growth
- extrapolation
- lurking variables
- causation
- common response
- confounding
- marginal distributions
- conditional distributions
- Simpson's paradox

Calculator Skills:



- LOG
- LinReg(a + bx)
- LinReg(ax + b)

4.1 Modeling Nonlinear Data (pp.176-206)

1. State the addition rule for logarithms. Give an example.

2. State the subtraction rule for logarithms. Give an example.

3. State the power rule for logarithms. Give an example.

4. Explain the difference between *linear growth* and *exponential growth*.

5. If the graph of the ordered pairs (x, y) is exponential, what type of graph is (x, logy)?

6. If the graph of the ordered pairs (x, y) is exponential, what type of graph is (logx, logy)?

4.2 Interpreting Correlation and Regression (pp.206-214)

1. What is *extrapolation*?

2. Define *lurking variable*.

3. If two variables have a strong positive association, then as one variable increases, the other variable also increases. Is it fair to say that an increase in one variable *causes* an increase in the other variable? Explain.

4. Define *causation*. Give an example.

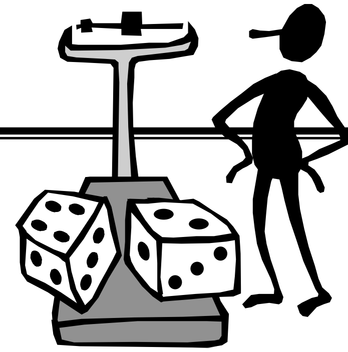
5. Define *common response*. Give an example.

6. Define *confounding*. Give an example.

4.3 Relations in Categorical Data (pp.215-229)

1.

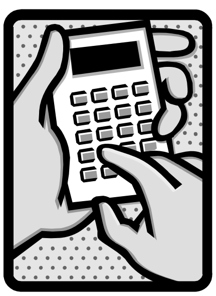
Chapter 5: Producing Data



Key Vocabulary:

- voluntary response sample
- confounded
- population
- sample
- design
- convenience sampling
- biased
- simple random sample
- table of random digits
- probability sample
- stratified random sample
- strata
- undercoverage
- nonresponse
- response bias
- sampling frame
- systematic random sample
- observational study
- experimental units
- subjects
- treatment
- factor
- level
- placebo effect
- control group
- randomization
- completely randomized experiment
- statistically significant
- replication
- hidden bias
- double-blind experiment
- block design
- matched pairs design
- simulation
- trial

Calculator Skills:



- STO
- ≤
- ≥
- and

5.1 Designing Samples (pp.245-264)

1. Why are *voluntary response samples* unreliable?

2. Explain the difference between a *population* and a *sample*?

3. Why might *convenience sampling* be unreliable?

4. What is a *biased* study?

5. What is meant by the *design* of a sample?

6. Define *simple random sample*.

7. What two properties of a *table of random digits* make it a good choice for creating a simple random sample?

8. What is a *stratified random sample*?

9. Give an example of *undercoverage* in a sample.

10. Give an example of *response bias* in a sample.

11. How can the wording of questions cause *bias* in a sample?

5.2 Designing Experiments (pp.265-285)

1. Explain the difference between an *observational study* and an *experiment*.

2. Explain the difference between *experimental units* and *subjects*.

3. Define *treatment*.

4. Give an example of at least two *levels* of a *factor* in an experiment.

5. Describe the *placebo effect*.

6. What is the significance of using a *control group*?

7. Define *randomization*.

8. Define *statistically significant*.

9. What are the advantages of a *double-blind study*?

10. Describe a *block design*.

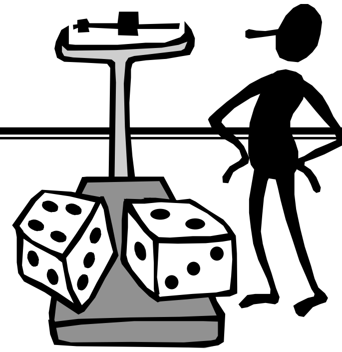
11. Describe a *matched pairs design*.

5.3 Simulating Experiments (pp.286-298)

1. What is *simulation*?

2. List the five steps for conducting a *simulation*:

Chapter 6: Probability



Key Vocabulary:

- trial
- random
- probability
- independence
- random phenomenon
- sample space
- S = {H, T}
- tree diagram
- replacement
- event
- P(A)
- Complement A^C
- disjoint
- Venn Diagram
- union (or)
- intersection (and)
- joint event
- joint probability
- conditional probability

6.1 Randomness (pp.310-317)

1. In statistics, what is meant by the term *random*?

2. In statistics, what is meant by *probability*?

3. What is *probability theory*?

4. In statistics, what is meant by an *independent* trial?

6.2 Probability Models (pp.317-340)

1. In statistics, what is a *sample space*?

2. In statistics, what is an *event*?

3. Explain why the probability of any *event* is a number between 0 and 1.

4. What is the sum of the probabilities of all possible *outcomes*?

5. Describe the probability that an *event* does not occur?

6. What is meant by the *complement* of an event?

7. When are two events considered *disjoint*?

8. What is the probability of two *disjoint* events?

9. Explain why the probability of getting heads when flipping a coin is 50%.

10. What is the *Multiplication Rule* for *independent* events?

11. Can *disjoint* events be *independent*?

12. If two events A and B are *independent*, what must be true about A^c and B^c?

6.3 More About Probability (pp.341-358)

1. What is meant by the *union* of two or more events? Draw a diagram.

2. State the addition rule for *disjoint* events.

3. State the general addition rule for *unions* of two events.

4. Explain the difference between the rules in #2 and #3.

5. What is meant by *joint probability*?

6. What is meant by *conditional probability*?

7. State the general multiplication rule.

8. How is the general multiplication rule different than the multiplication rule for independent events?

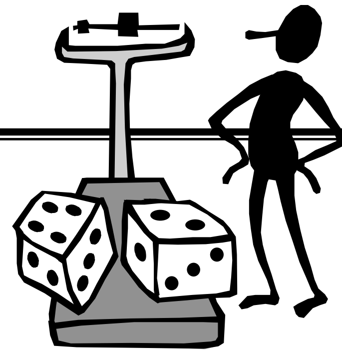
9. State the formula for finding conditional probability.

10. What is meant by the *intersection* of two or more events? Draw a diagram.

11. Explain the difference between the *union* and the *intersection* of two or more events.

12. State the formula used to determine if two events are *independent*.

Chapter 7: Random Variables



Key Vocabulary:

- random variable
- discrete random variable
- probability distribution
- probability histogram
- density curve
- probability density curve
- continuous random variable
- uniform distribution
- normal distribution
- μ_X
- μ_Y
- expected value
- Law of Large Numbers
- variance
- standard deviation

7.1 Discrete and Continuous Random Variables (pp.367-379)

1. What is a *discrete random variable*?

2. If *X* is a *discrete random variable*, what information does the *probability distribution of X* give?

3. In a *probability histogram* what does the height of each bar represent?

4. In a *probability histogram* what is the sum of the height of each bar?

5. What is a *continuous random variable*?

6. If *X* is a *discrete random variable*, how is the *probability distribution of X* described?

7. What is the area under a *probability density curve* equal to?

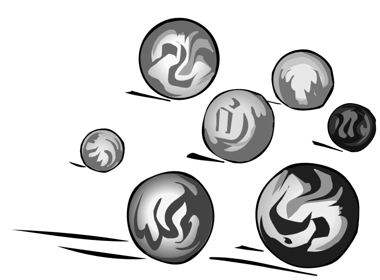
8. What is the difference between a *discrete random variable* and a *continuous random variable*?

9. If X is a *discrete random variable*, do $P(X > 2)$ and $P(X \geq 2)$ have the same value? Explain.

10. If X is a *continuous random variable*, do $P(X > 2)$ and $P(X \geq 2)$ have the same value? Explain.

11. How is a *normal distribution* related to *probability distribution*?

12. If a *normal distribution* is always a *probability distribution*, is a *probability distribution* always a *normal distribution*?



7.2 Means and Variances of Random Variables (pp.385-402)

1. Explain the difference between the notations $\bar{x}$ and μ_X .

2. What is meant by the *expected value* of X ?

3. How do you calculate the mean of a *discrete random variable X* ?

4. Explain the *Law of Large Numbers*.

5. Suppose $\mu_X = 5$ and $\mu_Y = 10$. According to the rules for means, what is μ_{X+Y} ?

6. Suppose $\mu_X = 2$. According to the rules for means, what is μ_{3+4X} ?

7. Explain how to calculate the *variance* of a *discrete random variable X* using the formula

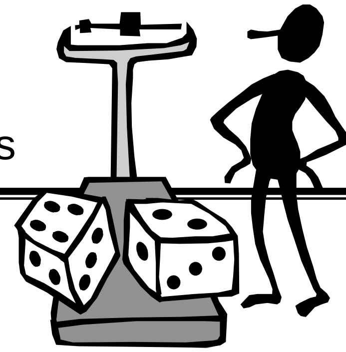
   $\sigma_X^2 = \sum (x_i - \mu_X)^2 p_i$.

8. Given the *variance* of a *random variable*, explain how to calculate the *standard deviation.*

9. Suppose $\sigma_X^2 = 2$ and $\sigma_Y^2 = 3$ and X and Y are independent random variables. According to the rules for variances, what is σ_{X+Y}^2 ? What is σ_{X+Y} ?

10. Suppose $\sigma_X^2 = 4$. According to the rules for variances, what is σ_{3+2X}^2 ? What is σ_{3+2X} ?

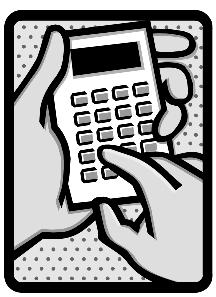
Chapter 8: The Binomial and Geometric Distributions



Key Vocabulary:

- binomial setting
- binomial random variable
- binomial distribution
- B(n, p)
- probability distribution function
- cumulative distribution function
- binomial coefficient
- $\binom{n}{k} = \frac{n!}{k!(n-1)!}$
- "n choose k"
- factorial
- geometric distribution

Calculator Skills:



- binompdf (n, p, X)
- binomcdf (n, p, X)
- randBin (n, p, #trials)
- geometpdf (p, # obs for success)
- geometcdf (p, # obs for success)

8.1 The Binomial Distributions (pp.415-434)

1. What are the four conditions for the *binomial setting*?

2. In the *binomial distribution*, what do parameters *n* and *p* represent?

3. What is meant by *B(n, p)* ?

4. What is the difference between a *probability distribution function* and a *cumulative distribution function*?

5. In the formula $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, what does n represent? What does k represent? What does the value of $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ represent?

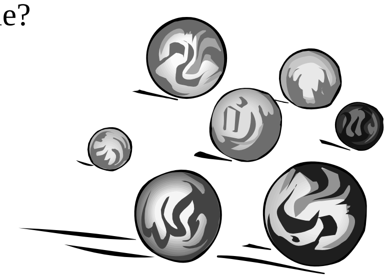
6. Complete the following table of values:

1!	1	1
2!	2 x 1	2
3!	3 x 2 x 1	6
4!	4 x 3 x 2 x 1	24

5!	5 x 4 x 3 x 2 x 1	
6!		
7!		
n!		

7. What is the value of $\frac{n!}{(n-1)!}$?

8. What are the mean and standard deviation of a binomial random variable?



8.2 The Geometric Distributions (pp. 434-444)

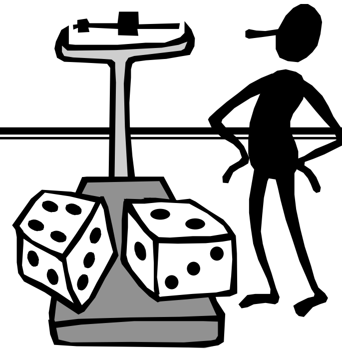
1. What are the four conditions for the *geometric setting*?

2. Explain the difference between the *binomial setting* and the *geometric setting*.

3. If X has a geometric distribution, what does $(1-p)^{n-1}p$ represent?

4. What is the *expected value* of a *geometric random variable*?

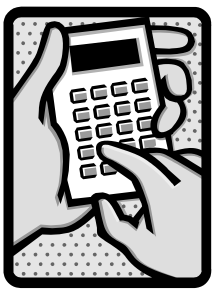
Chapter 9: Sampling Distributions



Key Vocabulary:

- parameter
- statistic
- sampling variability
- sampling distribution
- unbiased
- central limit theorem
- law of large numbers

Calculator Skills:



- randNorm(μ, σ, #trials)

9.1 Sampling Distributions (pp.456-469)

1. Explain the difference between a *parameter* and a *statistic*?

2. Explain the difference between p and $\hat{p}$?

3. What is *sampling variability*?

4. What is meant by the *sampling distribution* of a statistic?

5. When is a statistic considered *unbiased*?

6. How is the size of a sample related to the *spread* of the sampling distribution?

9.2 Sample Proportions (pp.472-479)

1. In an SRS of size *n*, what is true about the sampling distribution of $\hat{p}$ when the sample size *n* increases?

2. In an SRS of size *n*, what is the mean of the sampling distribution of $\hat{p}$?

3. In an SRS of size *n*, what is the standard deviation of the sampling distribution of $\hat{p}$?

4. What happens to the standard deviation of $\hat{p}$ as the sample size *n* increases?

5. When does the formula $\sqrt{\frac{p(1-p)}{n}}$ apply to the standard deviation of $\hat{p}$?

6. When the sample size *n* is large, the sampling distribution of $\hat{p}$ is approximately normal. What test can you use to determine if the sample is large enough to assume that the sampling distribution is approximately normal?

9.3 Sample Means (pp.481-494)

1. The mean and standard deviation of a population are *parameters*.
   What symbols are used to represent these *parameters*?

2. The mean and standard deviation of a sample are *statistics*.
   What symbols are used to represent these *statistics*?

3. Because averages are less variable than individual outcomes, what is true about the standard deviation of the sampling distribution of $\overline{x}$?

4. What is the mean of the sampling distribution of $\overline{x}$, if $\overline{x}$ is the mean of an SRS of size *n* drawn from a large population with mean μ and standard deviation σ?

5. What is the standard deviation of the sampling distribution of $\overline{x}$, if $\overline{x}$ is the mean of an SRS of size *n* drawn from a large population with mean μ and standard deviation σ?

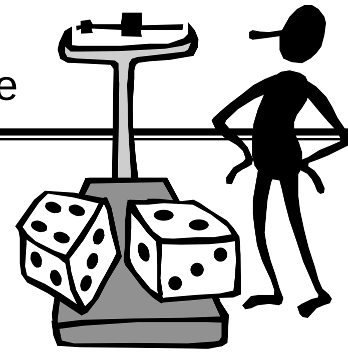
6. To cut the standard deviation of $\overline{x}$ in half, you must take a sample _____ times as large.

7. When should you use $\frac{\sigma}{\sqrt{n}}$ to calculate the standard deviation of $\overline{x}$?

8. What does the central limit theorem say about the shape of the sampling distribution of $\overline{x}$?

9. What is the law of large numbers?

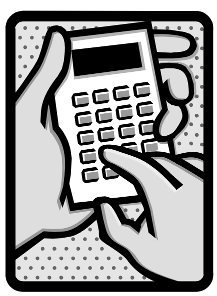
Chapter 10: Introduction to Inference



Key Vocabulary:

- confidence interval
- margin of error
- interval
- confidence level
- a level C confidence interval
- upper *p* critical value
- test of significance
- null hypothesis
- alternative hypothesis
- *p*-value
- statistically significant
- test statistic
- significance level
- *z* test statistic
- Hawthorne effect
- Type I Error
- Type II Error
- acceptance sampling
- power (of a test)

Calculator Skills:



- ZInterval
- Z-Test

10.1 Estimating with Confidence (pp.506-528)

1. In statistics, what is meant by a 95% *confidence interval*?

2. Sketch and label a *95% confidence interval* for the standard normal curve.

3. In a sampling distribution of $\bar{x}$, why is the interval of numbers between $\bar{x} \pm 2s$ called a *95% confidence interval*?

4. Define a *level C confidence interval*.

5. Sketch and label a *90% confidence interval* for the standard normal curve.

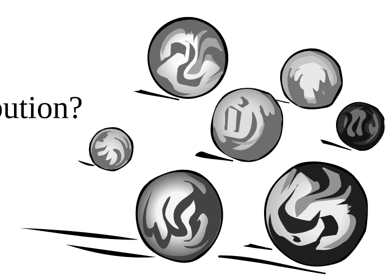
6. What does z^* represent?

7. What is the value of z^* for a *95% confidence interval*? Include a sketch.

8. What is the value of z^* for a *90% confidence interval*? Include a sketch.

9. What is the value of z^* for a *99% confidence interval*? Include a sketch.

10. What is meant by the *upper p critical value* of the standard normal distribution?



11. Explain how to find a *level C confidence interval* for an SRS of size *n* having unknown mean μ and known standard deviation σ.

12. What is meant by a *margin of error*?

13. Why is it best to have high *confidence* and a small *margin of error*?

14. What happens to the margin of error as *z** decreases? Does this result in a higher or lower confidence level?

15. What happens to the *margin of error* as σ decreases?

16. What happens to the *margin of error* as *n* increases? By how many times must the sample size *n* increase in order to cut the *margin of error* in half?

17. The formula used to determine the sample size *n* that will yield a confidence interval for a population mean with a specified margin of error *m* is $z^*\frac{\sigma}{\sqrt{n}} \leq m$. Solve for *n*.

10.2 Tests of Significance (pp.531-542)

1. What is a *null hypothesis*?

2. What is an *alternative hypothesis*?

3. In statistics, what is meant by the *P-value*?

4. If a *P-value* is small, what do we conclude about the *null hypothesis*?

5. If a *P-value* is large, what do we conclude about the *null hypothesis*?

6. How small should the *P-value* be in order to claim that a result is *statistically significant*?

7. Explain the difference between a *one-sided alternative hypothesis* and a *two-sided alternative hypothesis*.

8. What does a *test statistic* estimate?

9. What is meant by a *significance level*?

10.3 Using Significance Tests (pp.560-566)
10.4 Inference as Decision (pp. 567-577)

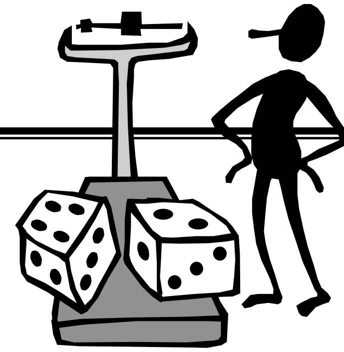
1. Significance tests are not always valid.
   What are some factors that can invalidate a test?

2. Explain the difference between a *Type I Error* and a *Type II Error*.

3. What is the relationship between the *significance level* α and the probability of *Type I Error*?

4. Describe how to calculate the power of a significance test.

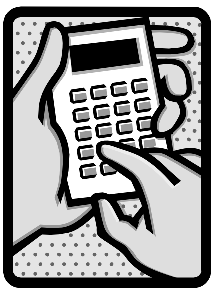
Chapter 11: Inference for Distributions



Key Vocabulary:

- standard error
- t distribution
- degrees of freedom
- t(k)
- z statistic
- one-sample t statistic
- two-sample t statistic
- robust
- power
- pooled

Calculator Skills:



- normalpdf (X)
- tpdf (X, df)
- ShadeNorm (leftendpoint, rightendpoint)
- Shade_t (leftendpoint, rightendpoint, df)
- TInterval
- T-Test
- 2-SampTTest
- 2-SampTInt

11.1 Inference for the Mean of a Population (pp.586-611)

1. Under what assumptions is *s* a reasonable estimate of σ?

2. In general, what is meant by the *standard error* of a statistic?

3. What is the *standard deviation* of the sample mean $\overline{x}$?

4. What is the *standard error* of the sample mean $\overline{x}$?

5. Describe the similarities between a *standard normal distribution* and a *t distribution*.

6. Describe the differences between a *standard normal distribution* and a *t distribution*.

7. How do you calculate the *degrees of freedom* for a *t distribution*?

8. What happens to the *t distribution* as the *degrees of freedom* increase?

9. How would you construct a level C confidence interval for μ if σ is unknown?

10. The *z*-Table gives the area under the standard normal curve to the left of *z*. What does the *t*-Table give?

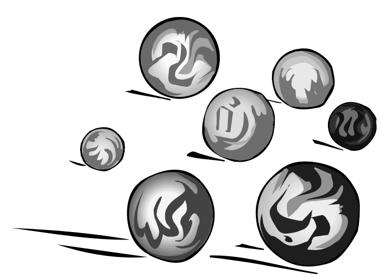
11. In a matched pairs *t procedure*, what is μ, the parameter of interest?

12. Samples from normal distributions have very few outliers. If your data contains outliers, what does this suggest?

13. If the size of the SRS is less than 15, when can we use *t procedures* on the data?

14. If the size of the SRS is at least 15, when can we use *t procedures* on the data?

15. If the size of the SRS is at least 40, when can we use *t procedures* on the data?



11.2 Comparing Two Means (pp.617-639)

1. How are two-sample problems different than one-sample problems?

2. Describe two different types of two-sample problems.

3. In a two-sample problem, what assumptions must be made for comparing two means?

4. In a two-sample problem, must/should the two sample sizes be equal?

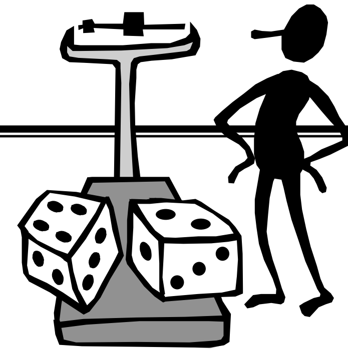
5. In a two-sample problem, what is the null hypothesis for comparing two means?

6. Explain how to standardize $\bar{x}_1 - \bar{x}_2$ if σ_1 and σ_2 are unknown.

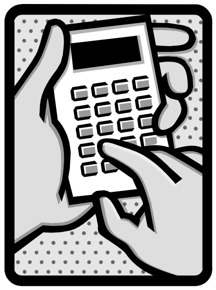
7. What assumption must you check if the sample sizes are small? How would you check?

8. If the two sample distributions for a two-sample problem are clearly skewed, how large should the samples be in order to use t procedures?

Chapter 12: Inference for Proportions



Calculator Skills:



- 1-PropZTest
- 1-PropZInt
- 2-PropZTest
- 2-PropZInt

12.1 Inference for a Population Proportion (pp.660-674)

1. In statistics, what is meant by a *sample proportion*?

2. Give the mean and standard deviation for the sampling distribution of $\hat{p}$?

3. How do you calculate the standard error of $\hat{p}$?

4. What assumptions must be met in order to use *z procedures* for inference about a proportion?

5. Describe how to construct a level C confidence interval for a population proportion.

6. For a one-sample hypothesis test where $H_0 : p = p_0$, what is the z test statistic?

7. What formula is used to determine the sample size necessary for a given margin of error?

12.2 Comparing Two Proportions (pp.678-689)

1. Give the mean and standard deviation for the sampling model of $\hat{p}_1 - \hat{p}_2$.

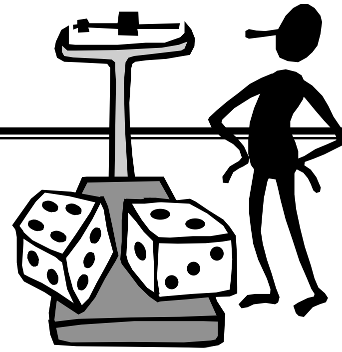
2. How do you calculate the standard error of $\hat{p}_1 - \hat{p}_2$?

3. What assumptions must be met in order to use *z procedures* for inference about two proportions?

4. Describe how to construct a level C confidence interval for the difference between two proportions, $p_1 - p_2$.

5. For a two-sample hypothesis test where $H_0 : p_1 = p_2$, what is the z test statistic?

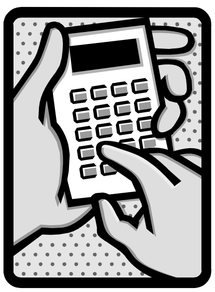
Chapter 13: Inference for Tables



Key Vocabulary:

- chi-square test for goodness of fit
- segmented bar chart
- chi-square statistic
- expected count
- observed count
- degrees of freedom
- chi-square distribution
- components of chi-square
- cell counts
- r x c table
- cell

Calculator Skills:



- sum ( )
- χ^2cdf (leftbound, rightbound, df)
- χ^2pdf (X, df)
- Shade χ^2 (leftbound, rightbound, df)
- χ^2-Test

13.1 Test for Goodness of Fit (pp.702-715)

1. What information does a *segmented bar chart* show?

2. Explain how to construct a *segmented bar chart*. Draw a sketch.

3. What is the *chi-square statistic*?

4. What is the difference between the notation X^2 and χ^2?

5. How many degrees of freedom does the *chi-square distribution* have?

6. As the *chi-square statistic* increases, what happens to the P-value?

7. What is the domain of a *chi-square distribution*?

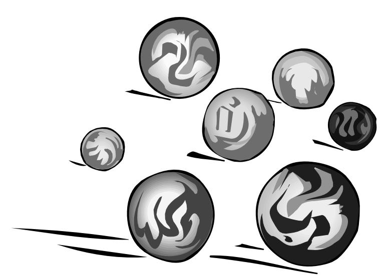
8. What is the shape of a *chi-square distribution*? What happens to the shape as the degrees of freedom increases?

9. State the null and alternative hypotheses for the *goodness of fit test*.

10. What conditions must be met in order to use the *goodness of fit test*?

11. What is meant by a *component* of chi-square?

12. What does the largest *component* of chi-square signify?



13.2 Inference for Two-Way Tables (pp.717-735)

1. Why is it necessary to perform follow-up analysis to a chi-square test?

2. What information is contained in a two-way table for a chi-square test?

3. State the null and alternative hypotheses for comparing more than two proportions.

4. How do you calculate the expected count in any cell of a two-way table when the null hypothesis is true?

5. How many degrees of freedom does a chi-square test for a two-way table with r rows and c columns have?

6. If you have an entire population, or a single SRS, with each individual classified according to both of two categorical variables, what is the null hypothesis for a chi-square test?