# TI 83/84 Calculator – The Basics of Statistical Functions

| **What you want to do >>>** | **Put Data in Lists** | **Get Descriptive Statistics** | **Create a histogram, boxplot, scatterplot, etc.** | **Find normal or binomial probabilities** | **Confidence Intervals or Hypothesis Tests** |
|---|---|---|---|---|---|
| How to start | STAT > EDIT > 1: EDIT ENTER | [after putting data in a list]<br>STAT > CALC ><br>1: 1-Var Stats ENTER | [after putting data in a list]<br>2nd STAT PLOT 1:Plot 1 ENTER | 2nd VARS | STAT > TESTS |
| What to do next | Clear numbers already in a list: Arrow up to L1, then hit CLEAR, ENTER.<br><br>Then just type the numbers into the appropriate list (L1, L2, etc.) | The screen shows:<br>1-Var Stats<br>You type:<br>2nd L1 or<br>2nd L2, etc. ENTER<br>The calculator will tell you $\bar{x}$, s, 5-number summary (min, Q1, med, Q3, max), etc. | 1. Select "On," ENTER<br>2. Select the type of chart you want, ENTER<br>3. Make sure the correct lists are selected<br>4. ZOOM 9<br>The calculator will display your chart | For **normal** probability, scroll to either<br>2: normalcdf(,<br>then enter low value, high value, mean, standard deviation; **or**<br>3:invNorm(, then enter area to left, mean, standard deviation.<br>For **binomial** probability, scroll to either 0:binompdf(, or A:binomcdf( , then enter n,p,x. | **Hypothesis Test**:<br>*Scroll to one of the following:*<br>1:Z-Test<br>2:T-Test<br>3:2-SampZTest<br>4:2-SampTTest<br>5:1-PropZTest<br>6:2-PropZTest<br>C:$X^2$-Test<br>D:2-SampFTest<br>E:LinRegTTest<br>F:ANOVA(<br>**Confidence Interval:**<br>*Scroll to one of the following:*<br>7:ZInterval<br>8:TInterval<br>9:2-SampZInt<br>0:2-SampTInt<br>A:1-PropZInt<br>B:2-PropZIn |

Other points: (1) To **clear the screen**, hit 2nd, MODE, CLEAR
(2) To enter a **negative number**, use the negative sign at the bottom right, not the negative sign above the plus sign.
(3) To **convert a decimal to a fraction**: (a) type the decimal; (b) MATH > Frac ENTER

# Frank's Ten Commandments of Statistics

1. The probability of choosing one thing with a particular characteristic equals the percentage of things with that characteristic.

2. Samples have STATISTICS. Populations have PARAMETERS.

3. "Unusual" means *more than 2 standard deviations* away from the mean; "usual" means *within* 2 standard deviations of the mean.

4. "Or" means Addition Rule; "and" means Multiplication Rule

5. If Frank says Binomial, I say npx.

6. If σ (sigma/the standard deviation *of the population*) is **known**, use Z; if σ is **unknown**, use T.

7. In a Hypothesis Test, the claim is ALWAYS about the **population**.

8. In the Traditional Method, you are comparing POINTS (the Test Statistic and the Critical Value); in the P-Value Method, you are comparing AREAS (the P-Value and α (alpha)).

9. If the P-Value is less than α (alpha), reject $H_0$ ("If P is low, $H_0$ must go").

10. The Critical Value (point) sets the boundary for α (area). The Test Statistic (point) sets the boundary for the P-Value (area).

# Chapters 3-4-5 – Summary Notes

## Chapter 3 – Statistics for Describing, Exploring and Comparing Data

### Calculating Standard Deviation

$s = \sqrt{\dfrac{\sum(x-\bar{x})^2}{n-1}}$ Example:

| $x$ | $x-\bar{x}$ | $(x-\bar{x})^2$ |
|---|---|---|
| 1 | -5 | 25 |
| 3 | -3 | 9 |
| 14 | 8 | 64 |
| Total | | 98 |

$\bar{x}$ = 6 (18/3)

$s = \sqrt{\dfrac{98}{3-1}} = \sqrt{49} = 7$

### Finding the Mean and Standard Deviation from a *Frequency* Distribution

| Speed | Midpoint ($x$) | Frequency ($f$) | $x^2$ | $f \cdot x$ | $f \cdot x^2$ |
|---|---|---|---|---|---|
| 42-45 | 43.5 | 25 | 1892.25 | 1087.5 | 47306.25 |
| 46-49 | 47.5 | 14 | 2256.25 | 665 | 31587.50 |
| 50-53 | 51.5 | 7 | 2652.25 | 360.5 | 18565.75 |
| 54-57 | 55.5 | 3 | 3080.25 | 166.5 | 9240.75 |
| 58-61 | 59.5 | 1 | 3540.25 | 59.5 | 3540.25 |
| | | 50 | | 2339 | 110240.50 |

$\bar{x} = \dfrac{\sum(f \cdot x)}{\sum f}$ **, so** $\bar{x} = \dfrac{2339}{50} \approx$ **46.8**

$s = \sqrt{\dfrac{n[\sum(f \cdot x^2)] - [\sum(f \cdot x)]^2}{n(n-1)}}$**, so s** $= \sqrt{\dfrac{50[110240.5] - 2339^2}{50(49)}} = \sqrt{\dfrac{41104}{2450}} \approx \sqrt{16.78} \approx$ **4.1**

### Percentiles and Values

The percentile of value x =

$$\frac{number\ of\ values < x}{total\ number\ of\ values} \cdot 100$$

(round to nearest whole number)

To find the value of percentile $k$:

$L = \dfrac{k}{100} \cdot n$; this gives the location of the value we want; if it's not a whole number, we go up to the next number. If it is a whole number, then the answer is the mean of that number and the number above it.

*****Using the Calculator:** ***To find mean & standard deviation of a frequency distribution or a probability distribution***: **First**: $\boxed{\text{STAT}}$ > EDIT $\boxed{\text{ENTER}}$, then in L1 put in all the $x$ values (*midpoints* if it's a frequency distribution); in L2 put in frequencies or probabilities as applicable. **Second**: $\boxed{\text{STAT}}$ > CALC $\boxed{\text{ENTER}}$ 1-Var Stats $\boxed{\text{2ND}}$ L1, $\boxed{\text{2ND}}$ L2 $\boxed{\text{ENTER}}$. The screen shows the mean ($\bar{x}$) and the standard deviation, either Sx (if it's a *frequency* distribution) or σx (if it's a *probability* distribution).

## Chapter 4 - Probability

**Addition Rule ("OR")**
P(A or B) = P(A) + P(B) – P(A and B)

**Multiplication Rule ("AND")**
P(A and B) = P(A) · P(B|A)

**Conditional Probability**
P(B|A) = $\dfrac{P(A\ and\ B)}{P(A)}$

Find the probability of **"at least 1"** girl out of 3 kids, with boys and girls equally likely.

*"All boys" means #1 is a boy AND #2 is a boy AND #3 is a boy, so we use the Multiplication Rule:
.5 x .5 x .5 = .125

| | |
|---|---|
| 0 girls) = P(all boys) = **.125*** | P(at least 1 girl) = P(1, 2 or 3 girls) = 1 minus .125 = **.875** |

These are **complements**, so their combined probability must = 1.

**Fundamental Counting Rule**: For a sequence of two events in which the first event can occur $m$ ways and the second event can occur $n$ ways, the events together can occur a total of $m \cdot n$ ways.
**Factorial Rule**: A collection of $n$ different items can be arranged in order $n!$ different ways. (Calculator Example: To get 4!, hit 4$\boxed{\text{MATH}}$>PRB>4$\boxed{\text{ENTER}}$

**Permutations Rule (Items all Different)**
1. $n$ different items available.
2. Select $r$ items without replacement
3. Rearrangements of the same items are considered to be different sequences (ABC is counted separately from CBA)

**Calculator** example: n = 10, r = 8, so ${}_{10}P_8$
Hit 10 $\boxed{\text{MATH}}$ > PRB > 2, then 8 $\boxed{\text{ENTER}}$ = 1814400

**Permutations Rule (Some Items Identical)**
1. $n$ different items available, and some are identical
2. Select all $n$ items without replacement
3. Rearrangements of distinct items are considered to be different sequences.

\# of permutations = $\dfrac{n!}{n_1!\, n_2! \cdots n_k!}$

**Combinations Rule**
1. $n$ different items available.
2. Select $r$ items without replacement
3. Rearrangements of the same items are considered to be the same sequence (ABC is counted the same as CBA)

**Calculator** example: n = 10, r = 8, so ${}_{10}C_8$
Hit 10 $\boxed{\text{MATH}}$ > PRB > 3, then 8 $\boxed{\text{ENTER}}$ = 45

## Formulas for Mean and Standard Deviation

| | All Sample Values | Frequency Distribution | Probability Distribution |
|---|---|---|---|
| **Mean** | $\bar{x} = \frac{\sum x}{n}$ | $\bar{x} = \frac{\sum(f \cdot x)}{\sum f}$ | $\mu = \sum(x \cdot P(x))$ |
| **Std Dev** | $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$ | $s = \sqrt{\frac{n[\sum(f \cdot x^2)] - [\sum(f \cdot x)]^2}{n(n-1)}}$ | $\sigma = \sqrt{\sum[x^2 \cdot P(x)] - \mu^2}$ |

## Chapter 5 - Discrete *Probability* Distributions

**Sec. 5.2**

A **random variable** is simply a number that can change, based on chance. It can either be **discrete** (*countable*, like how many eggs a hen might lay), or **continuous** (like how much a person weighs, which is *not* something you can *count*). Example: The number of Mexican-Americans in a jury of 12 members is a random variable; it can be anywhere between 0 and 12. And it is a discrete random variable, because it is a number you can count.

To find the mean and standard deviation of a probability distribution **by hand**, you need 5 columns of numbers: (1) $x$; (2) $P(x)$; (3) $x \cdot P(x)$; (4) $x^2$; (5) $x^2 \cdot P(x)$.
**Using the Calculator:** To find the mean and standard deviation of a probability distribution, **First**: $\boxed{\text{STAT}}$ > EDIT, then in L1 put in all the *x* values, and in L2 put in the probability for each *x* value. **Second**: $\boxed{\text{STAT}}$ > CALC > 1-Var Stats > 1-Var Stats L1, L2 $\boxed{\text{ENTER}}$.

**Sec. 5.3 – 5.4 – Binomial Probability**

| **Requirements** | **Formulas** | **Using the Calculator** |
|---|---|---|
| ___ **F**ixed number of trials<br>___ **I**ndependent trials<br>___ **T**wo possible outcomes<br>___ **Constant** probabilities | $\mu = n \cdot p$<br>$\sigma = \sqrt{npq}$<br>$q = 1 - p$ | 1. To get the probability of a **specific** number: 2nd $\boxed{\text{VARS}}$ binom**p**df (n, p, x) (which gives you the probability of getting **exactly** $x$ successes in $n$ trials, when $p$ is the probability of success in 1 trial).<br>2. To get a **cumulative** probability: 2nd $\boxed{\text{VARS}}$ binom**c**df ($n$, $p$, $x$) (which gives you the probability of getting **up to** $x$ successes in $n$ trials, when $p$ is the probability of success in 1 trial). *IMPORTANT: there are variations on this one, which we will talk about. Be sure to get them clear in your mind.*<br><br>At most/less than or equal to: ≤ binom**c**df($n$,$p$,$x$)<br>Less than: < binom**c**df($n$,$p$,$x$-1)<br>At least/greater than or equal to: ≥ 1 *minus* binom**c**df($n$,$p$,$x$-1)<br>Greater than/more than: > 1 *minus* binom**c**df($n$,$p$,$x$) |

| **Symbol Summary** | | **Sample** | **Population** |
|---|---|---|---|
| | How many? | $n$ | $N$ |
| | Mean | $\bar{x}$ | $\mu$ |
| | Proportion | $\hat{p}$ | $p$ |
| | Standard Deviation | $s$ | $\sigma$ |
| | Correlation Coefficient | $r$ | $\rho$ |

## Chapters 6-7-8 – Summary Notes

<table>
<tr><th>Ch</th><th>Topic</th><th>Calculator</th><th>Formulas, Tables, Etc.</th></tr>
<tr><td>6</td><td>Normal Probability Distributions<br>3 Kinds of problems:<br>1. You are given a point (value) and asked to find the corresponding area (probability)<br>1a. Central Limit Theorem. Just like #1, except $n > 1$.<br>2. You are given an area (probability) and asked to find the corresponding point (value).<br>3. Normal as approximation to binomial</td><td>1. $\boxed{2^{nd}}$ $\boxed{\text{VARS}}$ normalcdf (low, high, $\mu$,$\sigma$)<br>1a. $\boxed{2^{nd}}$ $\boxed{\text{VARS}}$ normalcdf (low, high, $\mu$ , $\sigma/\sqrt{\boldsymbol{n}}$)<br>2. $\boxed{2^{nd}}$ $\boxed{\text{VARS}}$ invNorm (area to left, $\mu$, $\sigma$)<br>3. Step 1: Using binomial formulas, find mean and standard deviation.</td><td>$$z = \frac{x - \mu}{\sigma}$$<br>Table A-2.<br>3. (cont'd – Normal as approximation to binomial) – Step 2:<br><u>If you are asked to find</u> — <u>Then in calculator</u><br>P(at least x) — normalcdf(x-.5,1E99,$\mu$,$\sigma$)<br>P(more than x) — normalcdf(x+.5,1E99,$\mu$,$\sigma$)<br>P(x or fewer) — normalcdf(-1E99,x+.5,$\mu$,$\sigma$)<br>P(less than x) — normalcdf(-1E99,x-.5,$\mu$,$\sigma$)</td></tr>
<tr><td>7</td><td>Confidence Intervals<br>1. Proportion<br>$\hat{p} - E < p < \hat{p} + E$<br>2. Mean (z or t?)<br>$\bar{x} - E < \mu < \bar{x} + E$<br>3. Standard Deviation<br>$\sqrt{\frac{(n-1)s^2}{X_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{X_L^2}}$</td><td>1. $\boxed{\text{STAT}}$ > TEST > 1PropZInt<br>Minimum Sample Size: $\boxed{\text{PRGM}}$ NPROP<br>2. $\boxed{\text{STAT}}$ > TEST > ZInt OR $\boxed{\text{STAT}}$ > TEST > TInt (use Z if σ is known, T if σ is unknown)<br>Minimum Sample Size: $\boxed{\text{PRGM}}$ NMEAN<br>3. $\boxed{\text{PRGM}}$ >INVCHISQ (to find $X_L^2$ and $X_R^2$)<br>$\boxed{\text{PRGM}}$ > CISDEV (to find Conf. Interval.)</td><td>1. $\hat{p}$ = sample proportion; $E = z_{\alpha/2}\sqrt{\hat{p}\hat{q}/n}$<br>Min. Sample Size ($\hat{p}$ unknown): $n = \frac{[z_{\alpha/2}]^2 \cdot .25}{E^2}$<br>Min. Sample Size ($\hat{p}$ known): $n = \frac{[z_{\alpha/2}]^2 \cdot \hat{p}\hat{q}}{E^2}$<br>2. $\bar{x}$ = sample mean; $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ (σ known)<br>or $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$ (σ unknown)<br>Min. Sample Size: $n = \left[\frac{z_{\alpha/2} \cdot \sigma}{E}\right]^2$<br>3. Use Table A-4 to find $X_L^2$ and $X_R^2$</td></tr>
<tr><td>8</td><td>Hypothesis Tests<br>1. Proportion<br>2. Mean (z or t?)<br>3. Standard Deviation<br>If P-Value < α, reject $H_0$; if P-Value > α fail to reject $H_0$.</td><td>1. $\boxed{\text{STAT}}$ > TEST > 1PropZTest<br>2. $\boxed{\text{STAT}}$ > TEST > ZTest OR TTest<br>3. $\boxed{\text{PRGM}}$ > TESTSDEV</td><td>1. Test Statistic: $z = \frac{\hat{p} - p}{\sqrt{pq/n}}$<br>2. Test Statistic: $z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma/\sqrt{n}}$ OR $t = \frac{\bar{x} - \mu_{\bar{x}}}{s/\sqrt{n}}$<br>3. Test Statistic: $X^2 = \frac{(n-1) \cdot s^2}{\sigma^2}$<br>See additional sheet on 1-sentence statement and finding Critical Value.</td></tr>
</table>

# Hypothesis Tests

| 1-Sentence Statement/Final Conclusion | | |
|---|---|---|
| | **Claim is $H_0$** | **Claim is $H_1$** |
| **Reject $H_0$ (Type 1 – reject true $H_o$)** | There is sufficient evidence to warrant rejection of the claim that . . . | The sample data support the claim that . . . |
| **Fail to reject $H_0$ (Type II – fail to reject false $H_o$)** | There is *not* sufficient evidence to warrant rejection of the claim that . . . | There is *not* sufficient sample evidence to support the claim that . . . |

**<u>Hypoth Test Checklist</u>**

___ CLAIM
___ HYPOTHESES
___ SAMPLE DATA
___ α
___ CALCULATOR: P-VALUE, TEST STATISTIC
___ CONCLUSIONS

**To find Critical Value**
**(required only for Traditional Method, not for P-Value Method)**

***Critical Z-Value***

| | | |
|---|---|---|
| **Left-Tail Test** *(1 negative CV)*<br>$\boxed{2^{nd}}$ $\boxed{\text{VARS}}$ invNorm(α) $\boxed{\text{ENTER}}$ | **Right-Tail Test** *(1 positive CV)*<br>$\boxed{2^{nd}}$ $\boxed{\text{VARS}}$ invNorm(1-α) $\boxed{\text{ENTER}}$ | **Two-Tail Test** *(1 neg & 1 pos CV)*<br>$\boxed{2^{nd}}$ $\boxed{\text{VARS}}$ invNorm(α/2) $\boxed{\text{ENTER}}$ |

***Critical T-Value (when you get to Chapter 9, for TWO samples, for DF use the <u>smaller</u> sample)***

| | | |
|---|---|---|
| **Left-Tail Test** *(1 negative CV)*<br>$\boxed{\text{PRGM}}$ > INVT $\boxed{\text{ENTER}}$<br>AREA FROM LEFT = α<br>DF = n-1 ; then hit $\boxed{\text{ENTER}}$ | **Right-Tail Test** *(1 positive CV)*<br>$\boxed{\text{PRGM}}$ > INVT $\boxed{\text{ENTER}}$<br>AREA FROM LEFT = 1-α<br>DF = n-1; then hit $\boxed{\text{ENTER}}$ | **Two-Tail Test** *(1 neg & 1 pos CV)*<br>$\boxed{\text{PRGM}}$ > INVT $\boxed{\text{ENTER}}$<br>AREA FROM LEFT = α/2<br>DF = n-1; then hit $\boxed{\text{ENTER}}$ |

***Critical $X^2$-Value***

| | | |
|---|---|---|
| **Left-Tail Test** *(1 positive CV)*<br>$\boxed{\text{PRGM}}$ > INVCHISQ $\boxed{\text{ENTER}}$ $\boxed{\text{ENTER}}$<br>DF = n-1 $\boxed{\text{ENTER}}$ $\boxed{\text{ENTER}}$<br>AREA TO RIGHT = 1 - α | **Right-Tail Test** *(1 positive CV)*<br>$\boxed{\text{PRGM}}$ > INVCHISQ $\boxed{\text{ENTER}}$ $\boxed{\text{ENTER}}$<br>DF = n-1 $\boxed{\text{ENTER}}$ $\boxed{\text{ENTER}}$<br>AREA TO RIGHT = α | **Two-Tail Test** *(MUST DO TWICE)*<br>$\boxed{\text{PRGM}}$ > INVCHISQ $\boxed{\text{ENTER}}$ $\boxed{\text{ENTER}}$<br>DF = n-1 $\boxed{\text{ENTER}}$ $\boxed{\text{ENTER}}$<br>AREA TO RIGHT: 1st time: α/2 ($X_R^2$)<br>2nd time: 1 - α/2 ( $X_L^2$) |

# Chapters 9-10-11 – Summary Notes

## Chapter 9 – Inferences from Two Samples (Be sure to use Hypothesis Test Checklist)

| | Proportions (9-2) | Means (9-3) (independent samples) | Matched Pairs (9-4) (dependent samples) |
|---|---|---|---|
| **Hypothesis Test (Be sure to use Hypoth Test checklist)** | Calculator: 2-PropZTest<br>Formulas:<br>$z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\bar{p}\bar{q}/n_1+\bar{p}\bar{q}/n_2}}$ *$p_1 - p_2$ always = 0.*<br>$\bar{p} = (x_1+x_2)/(n_1+n_2)$<br>$H_0$: $p_1 = p_2$; $H_1$: $p_1$ < or > or ≠ $p_2$ | Calculator: 2-SampleTTest<br>$t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{s_1^2/n_1+s_2^2/n_2}}$ *$\mu_1 - \mu_2$ always = 0*<br>$H_0$: $\mu_1 = \mu_2$; $H_1$: $\mu_1$ < or > or ≠ $\mu_2$ | Calculator: (1) Enter data in L1 and L2, L3 equals L1 – L2; (2) TTest (which gives you all the values you need to plug into the formula)<br>$t = \frac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$ *$\mu_d$ always = 0*<br>$H_0$: $\mu_d = 0$; $H_1$: $\mu_d$ < or > or ≠ 0 |
| **Confidence Interval** | Calculator: 2-PropZInt. | Calculator: 2-SampleTInt | Calculator: TInterval |
| | No formulas. If interval contains 0, then fail to reject. If $\alpha$ for a 1-tail Hypo. Test is .05, the CL for Conf. Int. is 0.9 (1 - 2$\alpha$). | | |

## Chapter 10 – Correlation and Regression (don't use formulas, just calculator; the important thing is interpreting results.)

| **Question to be answered** | **Calculator and Interpretation** (Anderson: show value of t but not formula) |
|---|---|
| Hypothesis Test: Is there a linear correlation between two variables, *x* and *y*? (10-2)<br>r is the sample correlation coefficient. It can be between -1 and 1. | **Calculator**: Enter data in L1 and L2, then STAT > Test > LinRegTTest. (Reg EQ > VARS, Y-VARS, Function, Y1). P-Value tells you if there *is* a linear correlation. The test statistic is r, which measures the *strength* of the linear correlation.<br>**Interpretation**: x is the **explanatory** variable; y is the **response** variable. $H_0$: there **no** linear correlation; $H_1$: there **is** a linear correlation. So, if P-Value **<** $\alpha$, you **reject $H_0$**, so there **IS** a linear correlation; if P-Value **>** $\alpha$, you **fail to reject $H_0$**, so there is **NO** linear correlation.<br>**Calculator**: To create a **scatterplot**: Enter data in L1 and L2, then LinRegTTest; then 2nd Y = Plot 1 On, select correct type of plot. Then Zoom 9 (ZoomStat). To delete regression line from graph, Y= , then clear equation from Y1. |
| When you have 2 variables *x* and *y*, how do you predict ***y*** when you are given a particular *x*-value? (10-3) | Two possible answers:<br>(1) *If there is a significant linear correlation*, then you need to determine the Regression Equation (y = a + bx; LinRegTTest gives you a and b), then just plug in the given x-value. Or the easy way to find *y* for any particular value of *x*:<br>**Calculator**: VARS Y-Vars 1:Function Enter Enter Input x-value in parentheses after Y1<br>(2) If there is ***no*** significant linear correlation, then the best predicted value for y = $\bar{y}$. **Calculator**: VARS, 5, 5, ENTER. |
| When you have 2 variables *x* and *y*, how do you predict an **interval estimate** for *y* when you are given a particular *x*-value? (10-4) | 1. PROGRAM, INVT ENTER Area from left is 1-$\alpha$/2, DF = n-2. This gives you t Critical Value.<br>2. PROGRAM, PREDINT ENTER Input t Critical Value from Step 1, then input X value given in the problem. Hit Enter *twice* to get the Interval. |
| How much of the **variation** in *y* is explained by the variation in *x*? (10-4) | The percentage of variation in y that is explained by variation in x is $r^2$, the coefficient of determination.<br>**Calculator**: to find $r^2$, enter data in L1 and L2, then LinRegTTest. It will give you $r^2$. 1 sentence conclusion: "[$r^2$]% of the variation in [y-variable in words] can be explained by the variation in [x-variable in words]."<br>**Total** Variation is $\sum(y-\bar{y})^2$. **Explained** Variation is Total Variation times $r^2$. **Unexplained** Variation is Total Variation minus Explained Variation. To find them all, put *x* values in L1; put *y* values in L2; LinRegTTest; then PRGM VARATION. |

**Chapter 10 - continued**

| | |
|---|---|
| Finding Total **Deviation**, Explained Deviation, and Unexplained Deviation, **for a point**. (10-4) | Variation and Deviation are similar, but different. Variation relates to ALL the points in a set of correlated data. Deviation relates to ONE specific point in a set of correlated data. **Total** Deviation for a specific point = $y - \bar{y}$ (the actual $y$-coordinate of the point *minus* the mean of all the $y$ values). **Explained** Deviation for the point = $\hat{y} - \bar{y}$ (the predicted value of $y$ when the $x$-coordinate of that point is plugged into the regression equation, *minus* the mean of all the $y$ values). **Unexplained** Deviation for the point = $y - \hat{y}$ (the actual $y$-coordinate of the point *minus* the predicted value of $y$ when the $x$-coordinate of that point is plugged into the regression equation). |

## Chapter 11 – Chi-Square ($X^2$) Problems (Hypothesis Tests, use checklist)

| **Claim to be tested** | **Calculator** | **Formulas, etc.** |
|---|---|---|
| **The claim that an observed proportion (*O*) < or = or > an expected proportion (*E*). This is called "goodness of fit." (11-2) *3 SAMPLES WITH PROPORTION*** | Enter O data in L1 and E data in L2. $\boxed{\text{PROGRAM}}$ BESTFIT $\boxed{\text{ENTER}}$ Enter number of categories (k), then $\boxed{\text{Enter}}$, Gives you Chi-Square ($X^2$) and P-Value | Two types of claims: *equal* proportions or *unequal* proportions:<br>• *Equal*: $H_0$: $p_1=p_2=p_3$; $H_1$: at least one is not equal<br>• *Unequal*: $H_0$: $p_1=.45$, $p_2 = .35$, $p_3 = .20$; $H_1$: at least one is not equal to the claimed proportion<br>Test Statistic is $X^2$ (Chi-Square) $X^2 = \sum \frac{(O-E)^2}{E}$<br>• Hypothesis Test is always right-tail<br>• To find P-Value from Test Stat: $X^2$cdf(TS,1E99,k-1) |
| **Given a table of data with rows and columns, the claim that the row variable is INDEPENDENT of the column variable. This is called "contingency tables." (11-3)** | $\boxed{2^{nd}}$ $\boxed{\text{MATRX}}$ EDIT, select [A], make sure number of rows and number of columns are correct, then enter the values in the matrix. $\boxed{\text{STAT}}$ Tests, $X^2$ – Test; Observed is [A] and Expected is [B], hit Calc; it gives you $X^2$ and P-Value. | Example hypotheses: $H_0$: pedestrian fatalities are *independent* of intoxication of driver ; $H_1$: pedestrian fatalities *depend* on intoxication of driver.<br>$X^2 = \sum \frac{(O-E)^2}{E}$<br>To find $X^2$ CV, df = (r-1)(c-1).<br>*E* (the expected value) for any cell = (row total x column total) / grand total, **OR** get all E-values on calculator with $\boxed{2^{nd}}$ $\boxed{\text{MATRX}}$ Edit [B] $\boxed{\text{Enter}}$ |

## Chapter 11 – Analysis of Variance (ANOVA) (Hypothesis Tests, use checklist)

| **Claim to be tested** | **Calculator** | **Formulas, etc. (must show formulas in both symbolic form and also with given data plugged in, but can get the answer from calculator)** |
|---|---|---|
| **The claim that 3 or more population means are all equal ($H_0$), or are not all equal ($H_1$). This is called "Analysis of Variance." (11-4) * 3 SAMPLES WITH MEAN*** | Enter data in L1, L2, etc. $\boxed{\text{STAT}}$ Tests, ANOVA (L1, L2, L3). Gives you F and PV, but also gives you Factor: df, SS and MS, and Error: df, SS and MS, which you need to plug into the formula. | The test statistic is F. $H_0$: $\mu_1 = \mu_2 = \mu_3$; $H_1$: at least one mean is not equal.<br>$F = \dfrac{\frac{\text{SS Factor}}{\text{df}}}{\frac{\text{SS Error}}{\text{df}}} = \dfrac{\text{MS Factor}}{\text{MS Error}}$<br>SS stands for "Sum of Squares"<br>MS stands for "Mean Squares"<br>MS(total)=SS(total)/(N-1). N=total number of values in all samples combined |

# Overview of Chapters 7, 8 and 9

**ONE-Sample Problems**

| Type of Problem | Chapter 7<br>Confidence Interval | Chapter 8<br>Hypothesis Test |
|---|---|---|
| Proportion | 1-Prop Z Int | 1-Prop Z Test |
| Mean – σ known | Z Interval | Z Test |
| Mean – σ unknown | T Interval | T Test |
| Standard Deviation | PRGM S2INT (Toma);<br>or<br>PRGM INVCHISQ + PRGM CISDEV | PRGM S2TEST (Toma);<br>or<br>PRGM TESTSDEV |

**TWO-Sample Problems**

| Type of Problem | Chapter 9<br>Confidence Interval | Chapter 9<br>Hypothesis Test |
|---|---|---|
| Proportion | 2-Prop Z Int | 2-Prop Z Test |
| Mean – Independent Samples | 2-Samp T Int | 2-Samp T Test |
| Mean – Dependent Samples (Matched Pairs) | T Interval | T Test |

**Minimum Sample Size Problems (Chapter 7)**

Chapter 7 also has another kind of problem, called **Minimum Sample Size** Problems.

These problems may ask you to find the minimum sample size, but usually they just say "How many . . .?"

Minimum Sample Size problems can be **Proportion** problems, **Mean** problems or **Standard Deviation** problems.

- **Proportion** problems: Use PRGM NPROP. This program asks you for SAMPLE P. If the problem gives you a Sample P (like 60%), then just enter that, as a decimal. If the problem does NOT give you a Sample P, then enter 0.5.
- **Mean** problems: Use PRGM NMEAN.

Both programs (NPROP and NMEAN) ask you for the Margin of Error. Sometimes the Margin of Error is clear from the problem. But if you see the word "within" in the problem, the margin of error is whatever comes immediately after the word "within."

- **Standard Deviation** problems: You are very rarely asked to find the Minimum Sample Size for a Standard Deviation problem. The only way to answer this kind of problem is by looking at the table in the Triola Text, which is on page 364 of the 4th Edition.

**Finding the key words in a problem (Ch. 6, 7 and 8)**

| Sample Problem | Key Words | Sample Statistics |
|---|---|---|
| 1. Assume that heights of men are **normally distributed**, with a mean of 69.0 in. and a standard deviation of 2.8 in. A day bed is 75 in. long. **Find the percentage** of men with heights that exceed the length of a day bed. | "normally distributed" tells you it's probably a Ch. 6 question.<br>"Find the percentage" tells you it's a **Type 1** Ch. 6 question (use **normalcdf**) | NA |
| 2. Assume that heights of men are **normally distributed**, with a mean of 69.0 in. and a standard deviation of 2.8 in. In designing a new bed, you want the length of the bed to equal or exceed the height of at least 95% of all men. **What is the minimum length** of this bed? | "normally distributed" tells you it's probably a Ch. 6 question.<br>"What is the minimum length" tells you it's **NOT** a Type 1 Ch. 6 question, so it's a Type 2 Ch. 6 question (use **invNorm**) | NA |
| 3. The cholesterol levels of men aged 18-24 are **normally distributed** with a mean of 178.1 and a standard deviation of 40.7. If 1 man aged 18-24 is randomly selected, **find the probability** that his cholesterol level is greater than 260. | "normally distributed" tells you it's probably a Ch. 6 question.<br>"Find the probability" tells you it's a **Type 1** Ch. 6 question (use normalcdf) | NA |
| 4. In a **poll** of 745 **randomly selected** adults, 589 said that it is morally wrong to not report all income on tax returns. **Construct a 95% confidence interval** estimate of the **percentage** of all adults who have that belief. | "poll" and "randomly selected" both tell you the sentence is about a **sample**.<br>"Construct a confidence interval" tells you you're constructing a confidence interval (Ch. 7), NOT testing a hypothesis (Ch. 8).<br>"Percentage" tells you it's a **proportion** problem, not a mean or standard deviation problem. | $n = 745$<br>$x = 589$ |
| 5. You must conduct a survey to determine the **mean** income reported on tax returns. **How many** randomly selected adults must you survey if you want to be 99% confident that the **mean** of the sample is **within** $500 of the true population mean? | "How many" tells you you are finding a **minimum sample size**.<br>"Mean" tells you you're finding a minimum sample size for a **mean**, not for a proportion.<br>"Within" tells you that the next thing you see ($500) is the **margin of error**. | NA |
| 6. A simple random sample of 37 weights of pennies has a mean of 2.4991g **and a standard deviation** of 0.0165g. **Construct a 99% confidence interval** estimate of the **mean** weight of all pennies. | "Construct a confidence interval" tells you you're constructing a confidence interval (Ch. 7), NOT testing a hypothesis (Ch. 8).<br>"Mean" tells you you're constructing a confidence interval for a **mean**, not for a proportion.<br>The phrase "and a standard deviation" is in a sentence about the **sample**, so you know the **sample** standard deviation (*s*), but *you don't know the **population** standard deviation (σ)*, so you will use **t**, not z. | $n = 37$<br>$\bar{x} = 2.4991$<br>$s = 0.0165$ |

| | | |
|---|---|---|
| 7. The Town of Newport has a new sheriff, who compiles **records** showing that among 30 recent robberies, the arrest **rate** is 30%, and she **claims** that her arrest **rate** is higher than the historical 25% arrest rate. **Test her claim**. | "Records" tells you you are getting information about a **sample**.<br>"Rate" tells you that this is a **proportion** problem, not a mean problem or standard deviation problem.<br>"Claim" and "Test her claim" tell you that this is a **Hypothesis Test** (Ch. 8) not a Confidence Interval (Ch. 7). | $n = 30$<br>$\hat{p} = .3$<br>$x = n \cdot \hat{p} = 9$ |
| 8. The totals of the individual weights of garbage discarded by 62 households in one week have a mean of 27.443 lb. **Assume** that the standard deviation of the weights is 12.458 lb. Use a 0.05 **significance level** to **test the claim** that the population of households as a **mean** less than 30 lb. | "Assume" tells you that the standard deviation you are being given is a **population** standard deviation (σ), so you will use **z**, not t.<br>"Test the claim" tell you that this is a **Hypothesis Test** (Ch. 8) not a Confidence Interval (Ch. 7).<br>"Mean" tells you you are testing a claim about a **mean**, not about a proportion or standard deviation.<br>"Significance level" is α (alpha) | $n = 62$<br>$\bar{x} = 27.443$<br>$\sigma = 12.458$<br><br>$\alpha = 0.05$ |

**Chapter 8 - Frank's Claim Buffet**

| The Claim in Words | Claim Buffet |
|---|---|
| 1. Test the claim that **<u>less than ¼</u>** of such adults smoke. | P / μ / σ | > / < / ≠ / = | Some number = **.25** |
| 2. Test the claim that **<u>most</u>** college students earn bachelor's degrees within 5 years | P / μ / σ | > / < / ≠ / = | Some number = **.5** | "most" just means more than half |
| 3. Test the claim that the **<u>mean</u>** weight of cars is **<u>less than 3700</u>** lb | P / μ / σ | > / < / ≠ / = | Some number = **3700** |
| 4. Test the claim that **<u>fewer than 20%</u>** of adults consumed herbs within the past 12 months. | P / μ / σ | > / < / ≠ / = | Some number = **.20** | 20% tells you the claim is about a proportion |

5. Test the claim that the thinner cans have a **<u>mean</u>** axial load that is **<u>less than 282.8</u>** lb

| | | |
|---|---|---|
| P | > | Some number = **282.8** |
| μ | < | |
| σ | ≠ | |
| | = | |

6. Test the claim that the sample comes from a population with a **<u>mean equal</u>** to **<u>74</u>**.

| | | |
|---|---|---|
| P | > | Some number = **74** |
| μ | < | |
| σ | ≠ | |
| | = | |

7. Test the claim that the **<u>standard deviation</u>** of the weights of cars is **<u>less than 520</u>**.

| | | |
|---|---|---|
| P | > | Some number = **520** |
| μ | < | |
| σ | ≠ | |
| | = | |