Visual Sample Plan Training Course
Version 2.0

Admin Building
Computer Training Room 16
Volpentest HAMMER Training & Education Center
2890 Horn Rapids Road
Richland, WA 99352
(509) 372-3143

http://www.hanford.gov/dqo/vsp_training/future_vsp.html

February 5, 2004

Session 1: 8:30 AM – 11:30 AM
Session 2: 1:00 PM – 4:00 PM

Presented by:

Sebastian Tindall,
Bechtel Hanford Inc.

Agenda

Visual Sample Plan Training Course

Version 2.0

Session 1: 8:30 AM – 11:30 AM
Session 2: 1:00 PM – 4:00 PM

Presented by:

Sebastian Tindall – Bechtel Hanford Inc., Richland, WA

Module 01: Agenda/Introductions
Module 02: Class Installation of VSP from PNNL VSP Web Site

- **Demonstration of INSTALL**
- **Class INSTALL**

Module 03: VSP Instruction Slide Exercise
Module 04: VSP Student Exercises
Module 05: VSP Student Exercise Solutions
Class Roster

02 - INSTALL Instructions for Visual Sample Plan 2.0

Instructions:

- Launch Internet Explorer
- Type http://www.hanford.gov/dqo and hit **Enter**
- Click on **"ERC VSP Training Course"** found on the left margin
- Click on **"Download/Install"**
- Click on **"Visual Sample Plan, Version 2.0 Installation"**
- Click on "**Down/Install VSP Software from PNNL**"
- Click on the hyperlink to **"Download** the Latest **Visual Sample Plan** version here**"**
- Enter, First Name, Last Name, company and e-mail address and click on the **"Submit"** button
- Click on "**Download Version 2.0 Final**"
- At the "File Download" dialog box, select **"Run this program from its current location"**
- VSP20.exe will be saved automatically to a temporary directory
- Select **"Yes"** to **"Do you want to install and run VSP20.exe from dqo.pnl.gov"?**
- On the **"Welcome to the InstallShield Wizard for Visual Sample Plan"** box, click on **Next**
- On the **License Agreement** box, select **"I accept the terms in the license agreement"**. Click on **Next**.
- On the **Destination Folder** box, click on **Next** (should install to the folder, c:\program files\visual sample plan)
- On the **Ready to Install the Program** box, click on **Install**.
- On the **InstallShield Wizard Completed** box, click on **Finish**. Visual Sample Plan should now be installed on your computer.
- Go to **Start - Programs - Visual Sample Plan** and Select **Visual Sample Plan.**

Note: The terms "view" and "window" are used interchangeably in Visual Sample Plan and in the instructions below.
For this exercise set: **OPTIONS→Preferences→Input LBGR/UBGR**

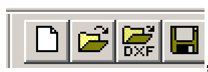
<u>Goal: To select the .dxf map file that you want to open.</u>

To make your screen look like Figure 1 below:

- From **Start→Programs→Visual Sample Plan**, open **Visual Sample Plan**.
- Click the **General (all inclusive) VSP** option in the **Select VSP Version** popup menu.
- Click the **Close** button at the bottom of the **VSP Advisor** help box.

- To load a drawing in the DXF file format, either:
  - From the main menu select **Map.**
  - Click **Load DXF…**, and then highlight **Millsite.dxf**.

Or

- Click the **Load Map** icon , and then highlight **Millsite.dxf**.

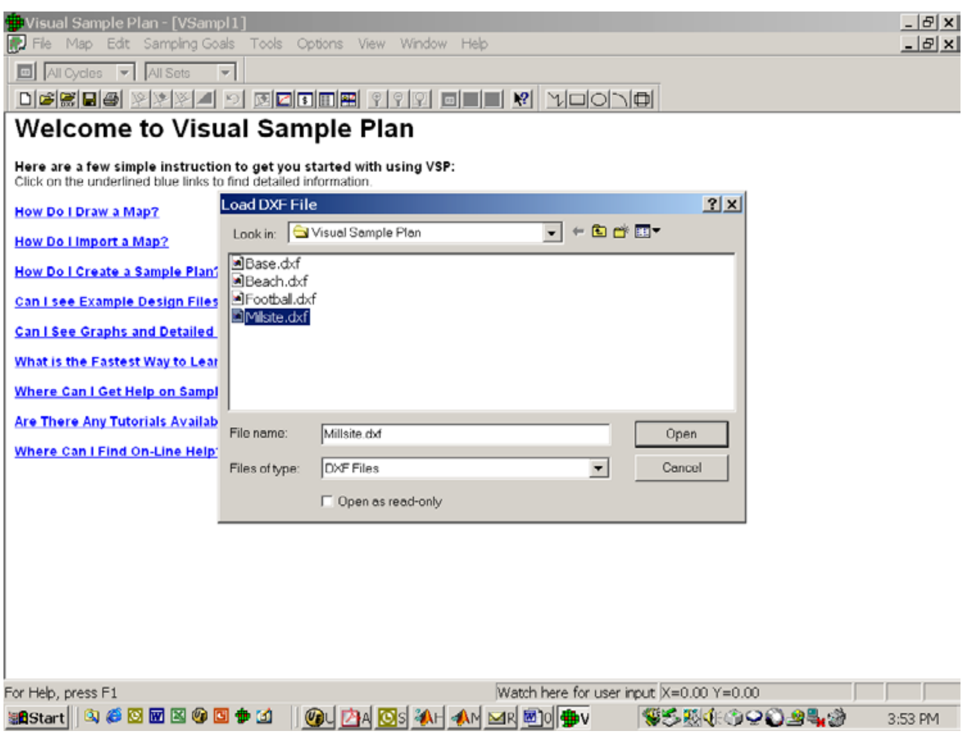
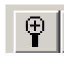
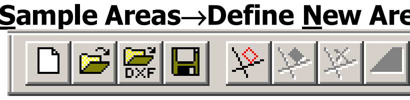


Figure 1. Dialog box for opening a DXF File

Presented by Bechtel Hanford, Inc.
Environmental Restoration Contractor

Goal: To open the selected .dxf map file and prepare to select a specific area within the map.

To make your screen look like Figure 2 below:

- With **Millsite.dxf** highlighted, click the **Open** button.
- Your Millsite map graphic will not fill the entire map window. To enlarge the map, select **View→Zoom In** from the main menu. Notice that the cursor now looks like a magnifying glass and the icon that looks like a magnifying glass with the plus sign is now selected. This is the **Zoom In** icon . Place the cursor on the oval (tailings pile) and click six or seven times to enlarge the map. Use the horizontal and vertical scroll bars to position your map like the one in Figure 2. Deselect **Zoom In** by clicking on the **Zoom In** icon. (Click the **Zoom Out** icon  if you need to shrink the map and then deselect it to return cursor to normal arrow.)
- From the main menu select **Edit→Sample Areas→Define New Area,**
- **Or** simply click the **New Area** icon .

------------------------▲

- A **Color** dialog box will open. The color yellow is selected by default; click **OK.**
- The cursor has changed to a crosshair and a yellow box of instructions appears.

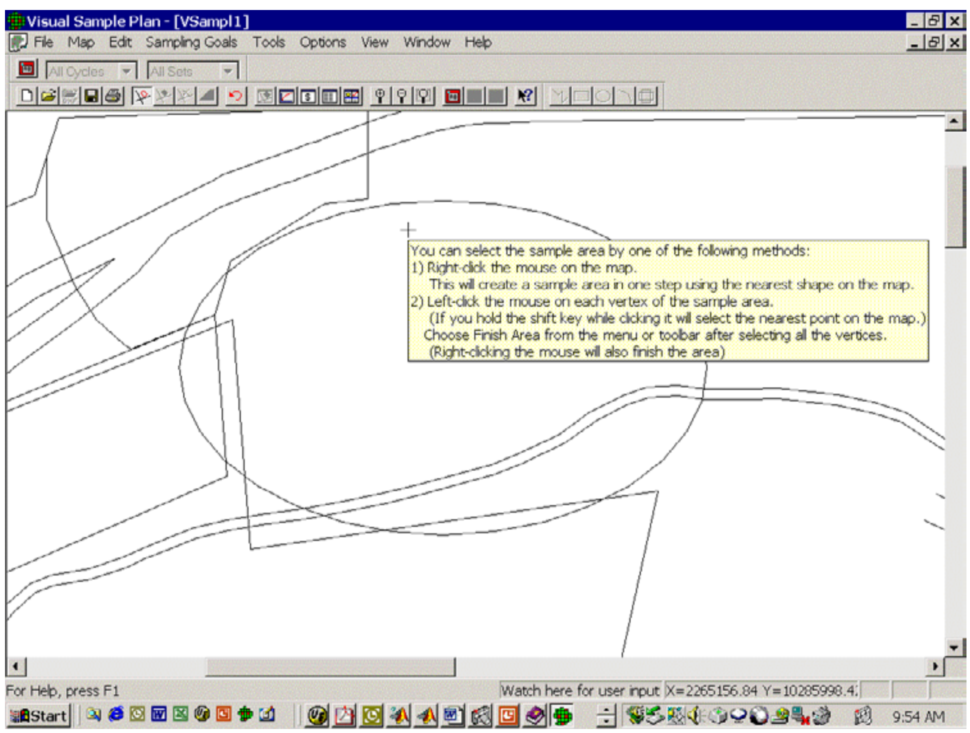


Figure 2. Tool tip for selecting a sample area

Goal: To finish selecting a specific area within the map.

To make your screen look like Figure 3 below:

- Move the crosshair to the upper right part of the oval shape near the top edge (away from the river) and right click with the mouse.
- A **Select area units** dialog box will open. Use the default area unit of **Feet^2** and click **OK**.
- The oval should be filled in with the color yellow.

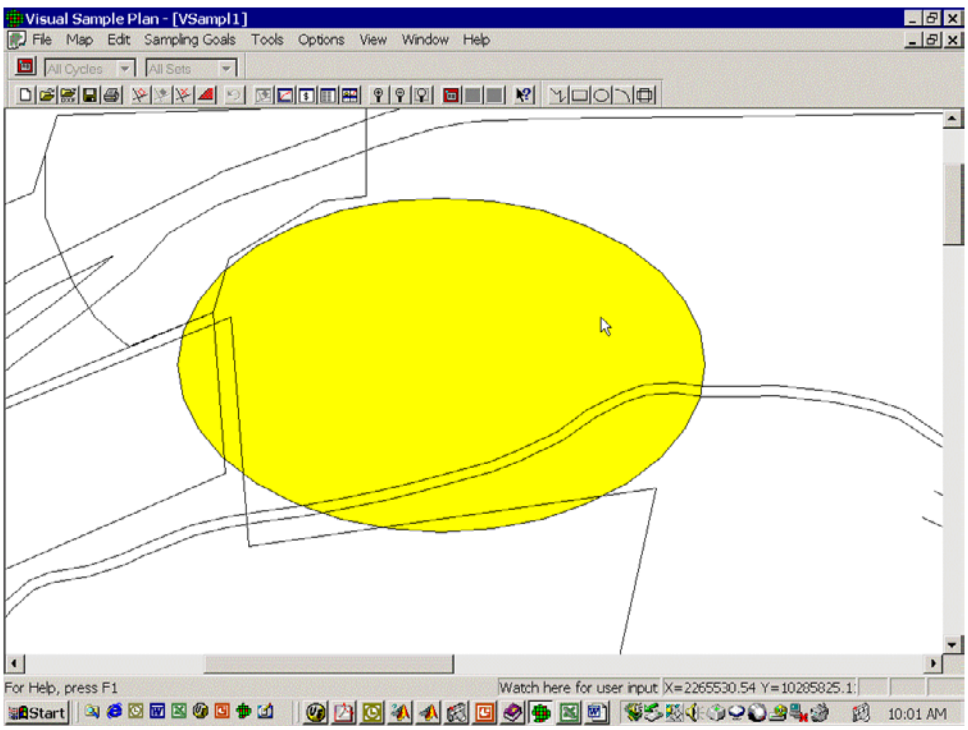
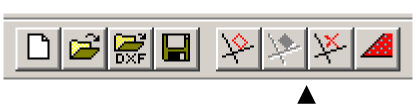


Figure 3. A sample area after it has been selected

NOTE: If you accidentally select the river or some other area instead of the oval, click on the **Remove Areas** icon and start over. It is the 7th icon from the left on the VSP toolbar.



Remove Areas icon

Presented by Bechtel Hanford, Inc.
Environmental Restoration Contractor

Goal: Create a sampling design to compare an average to an Action Level, data assumed normally distributed.

To make your screen look like Figure 4 below:

- From the main menu, select **Sampling Goals→Compare Average to Fixed Threshold→Can assume data will be normally distributed→Simple random sampling …**
- A **True Mean vs. Action Level** dialog box will appear. Enter the following values under the tab **One-Sample t-Test**:
- **Choose:** True Mean >= Action Level (Assume Site is Dirty).
- **False Rejection Rate (Alpha)** 1.0%
- **False Acceptance Rate (Beta)** 1.0%
- **Lower Bound of Gray Region** 8
- **Action Level** 10
- **Estimated Standard Deviation** 3

Note: Make sure that "Lower Bound of Gray Region" appears in the dialog box. If "Width of Gray Region" is showing instead, go to the Main Menu, Click **Options→Preferences→Input Delta** and change to "**Input LBGR**". Hereafter during these exercises, you can select either "Delta" or "LBGR". Just make sure that you make the necessary arithmetic adjustments.

Click **Apply.** Note the bottom line in the dialog box: **Minimum Samples in Survey Unit: 52**

- To make all sample points visible, move dialog box by placing mouse cursor in blue title bar, holding left mouse button down, and dragging box away from sample area.
- Note that the specific sample points on your map are likely to be in different locations. This is due to differences in the random locations for placing sampling points. If you click **Apply** repeatedly, you will see different random patterns of sampling locations.

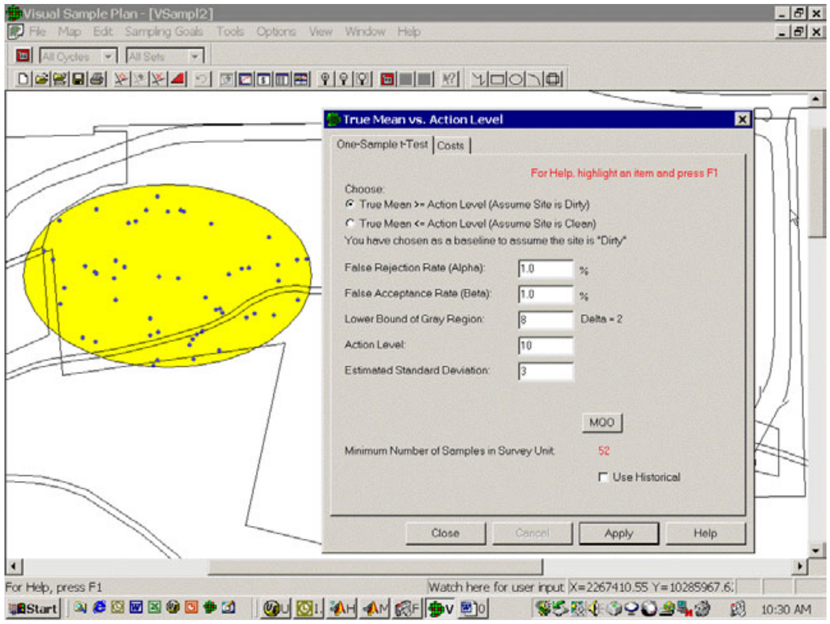


Figure 4. Dialog box for One-Sample t-Test

Goal: To simultaneously see the map view, the graph view, the report view and the coordinates view of the sampling design.

To make your screen look like Figure 5 below:

- Click the **Close** button to close the **True Mean vs. Action Level** dialog box.
- From the main menu, select **Window→Quad Window**.
- If necessary, use the horizontal and vertical scroll bars to move the yellow sampling area into the center of the visible map.

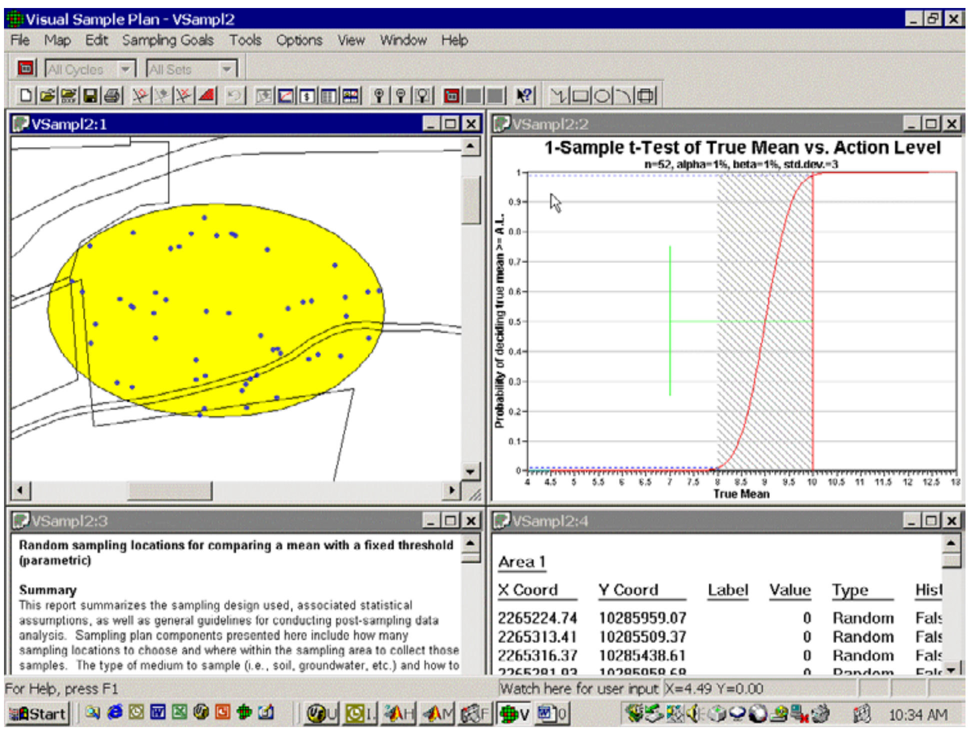


Figure 5. Example of the Quad Window

- If you choose, each window can be resized for better visibility. For example, place your cursor between the map window and the report window on the left side of the screen. The normal arrow cursor will change to a vertical double-headed arrow that can be used to change the size of a window. Please note that on slower PCs changing the size of the report window can take a few moments while the report is being updated.

Goal: To see how the probability of deciding that the site mean is above the action level changes as a function of the true mean of the site.

To make your screen look like Figure 6 below:

- From the **Quad Window** in Figure 5, enlarge the graph in the upper right hand corner until it covers the full screen (Click the **Maximize** button in the top right corner of the **Graph View**, i.e., a small square.)

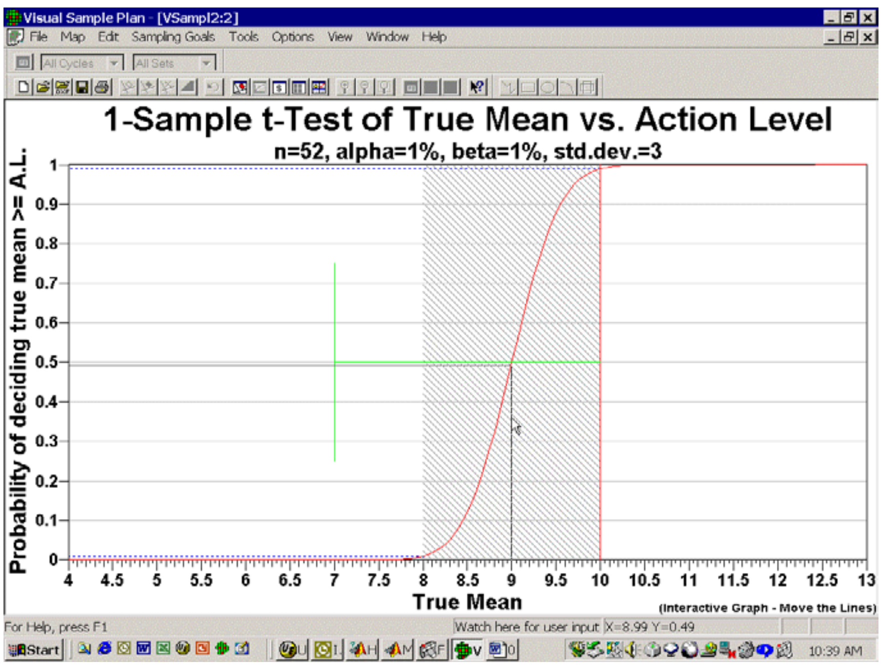


Figure 6. Example of the Graph View

The Graph View provides an interactive way to change the alpha error rate, beta error rate, the width of the gray region, etc. The dashed blue line near the top of the graph allows you to interactively change the alpha error rate. Similarly, the dashed blue line near the bottom of the graph allows you to change the beta error rate. The green vertical line allows you to change the estimated value of the standard deviation. Finally, dragging the left side of the gray region left or right changes its value.

Drag the lines representing these values and observe the changing results in the graph subtitle. For example, drag the top dashed blue line down so that the alpha error rate is reported as 5% in the subtitle. Now drag the bottom dashed blue line up so that the beta error rate is reported as 10%. Finally, drag the green line to the left until the std. dev. equals 4. Note the change in sample size from the original value of 52. (These interactive changes made on the graph do not change the values you entered in an open dialog box. However, if you close the dialog box and reopen it, the values from the graph will be the new defaults.)

Presented by Bechtel Hanford, Inc.
Environmental Restoration Contractor

Goal: Create a sampling design to compare an average to an Action Level, data not assumed normally distributed.

To make your screen look like Figure 7 below:

- From the main menu select **View→Map**.
- From the main menu, select **Sampling Goals→Compare Average to Fixed Threshold→Data not required to be normally distributed→Simple random sampling (Wilcoxon signed ranks test) …**
- A **True Mean or Median vs. Action Level** dialog box will appear. Be sure that it contains the following values under the **Wilcoxon Signed Rank Test** tab:
  - **Choose:** True Mean >= Action Level (Assume Site is Dirty).
  - **False Rejection Rate (Alpha)** 1.0%
  - **False Acceptance Rate (Beta)** 1.0%
  - **Lower Bound of Gray Region** 8
  - **Action Level** 10
  - **Estimated Standard Deviation** 3
  - If you click **Apply** repeatedly, you will see different random patterns of sampling locations.
  - Note the bottom line in the dialog box: **Minimum Number of Samples in Survey Unit: 60.** A parametric sampling design (e.g., the one-sample *t*-test) generally requires fewer samples than a nonparametric design (e.g., the Wilcoxon Signed Ranks test) to meet the same error tolerance rates, but a parametric design requires that we assume the sampling distribution of means (or the actual data distribution) is approximately normal.
  - EPA recommends using decision error rates of 1% as a starting point (EPA 2000a, p. 6-11). However, the DQO guidance acknowledges that "potential impacts on cost" may provide a rationale for relaxing these rates (EPA 2000a, p. 6-11). Make the following changes to the error rates and note the reduction in sample size:
  - **False Rejection Rate (Alpha)** 5.0%
  - **False Acceptance Rate (Beta)** 10.0%
  - Click **Apply**. The new sample size should now be: **24.** Whether this new sample size is justified depends on the consequences of decision errors and the relative costs of unnecessary cleanup vs. sampling costs.

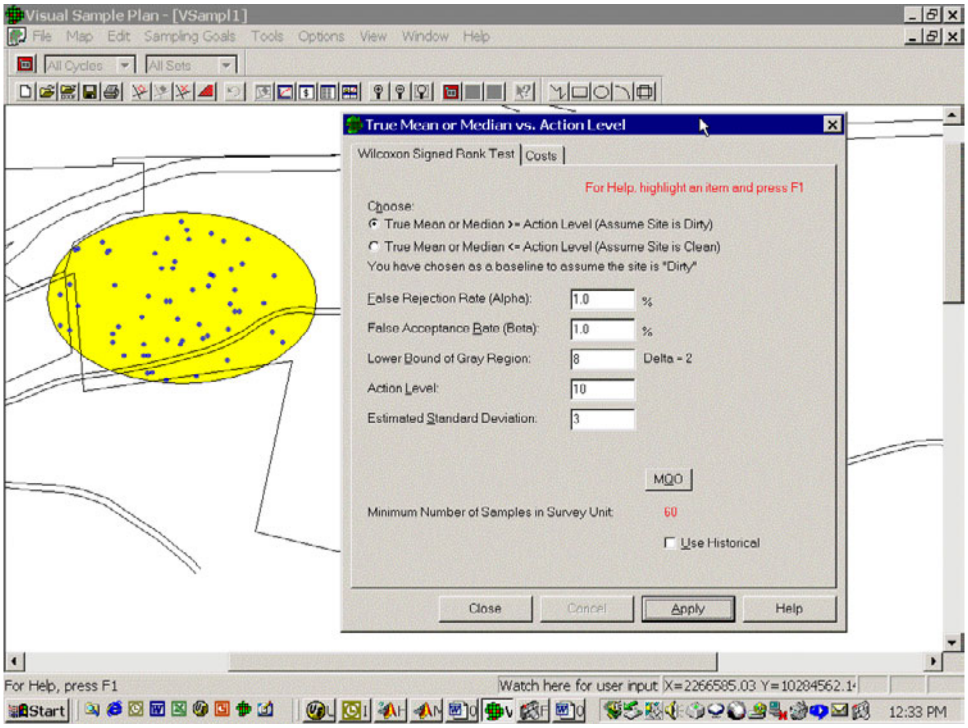


Figure 7. Wilcoxon Signed Rank Test dialog box and simple random sampling design

- **Close** the dialog box.
- From the main menu select **Window→Quad Window.**

Goal: Create a sampling design suitable for a Wilcoxon Signed Rank Test using systematic grid sampling.

To make your screen look like Figure 8 below:

- From the main menu select **Sampling Goals→Compare Average to Fixed Threshold→Data not required to be normally distributed→ Systematic grid sampling (Wilcoxon signed ranks test) ...**
- A **True Mean or Median vs. Action Level** dialog box will open. Verify that the last input values used in the previous example are the new defaults. If not, input the last input values noting that the alpha error = 5% and the beta error = 10%.
- Select the **Grid** tab.
- Select **Triangular** as the **Grid Type.**
- Check **Random Start**, if not already checked.
- Click **Apply.**
- Now position the dialog box so that the yellow sampling area is visible.
- A similar number of sampling locations are mapped as in Figure 7, but in a random-start grid pattern rather than a simple random pattern.

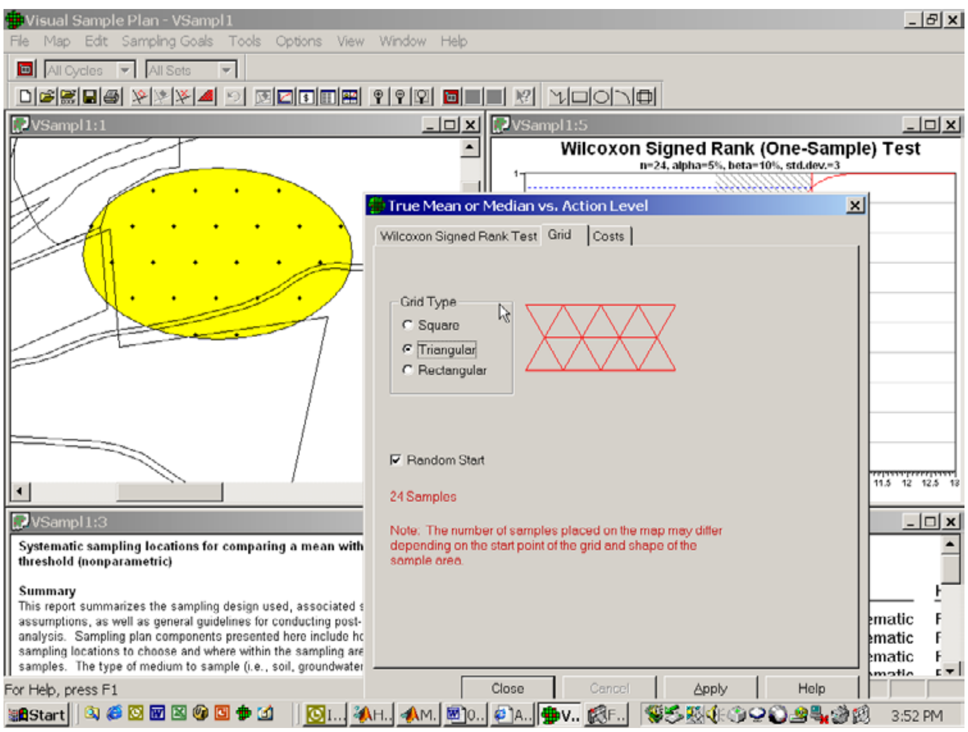


Figure 8. Example of random-start systematic grid and Wilcoxon Signed Rank Test dialog box

Goal: To learn how to change sampling costs for the Wilcoxon Signed Rank systematic grid sampling design above.

To make your screen look like Figure 9 below:

- Within the **True Mean or Median vs. Action Level** dialog box, select the **Costs** tab.
- Check that the following values are entered:
  - **Fixed Planning and Validation Cost $3000.00**
  - **Field Collection Cost per Sample $50.00**
  - **Measurement Cost per Analysis $500.00**
- Click **Apply.**
- The bottom-most line in the dialog box says **Total Cost for 24 Samples: $16200.00.**

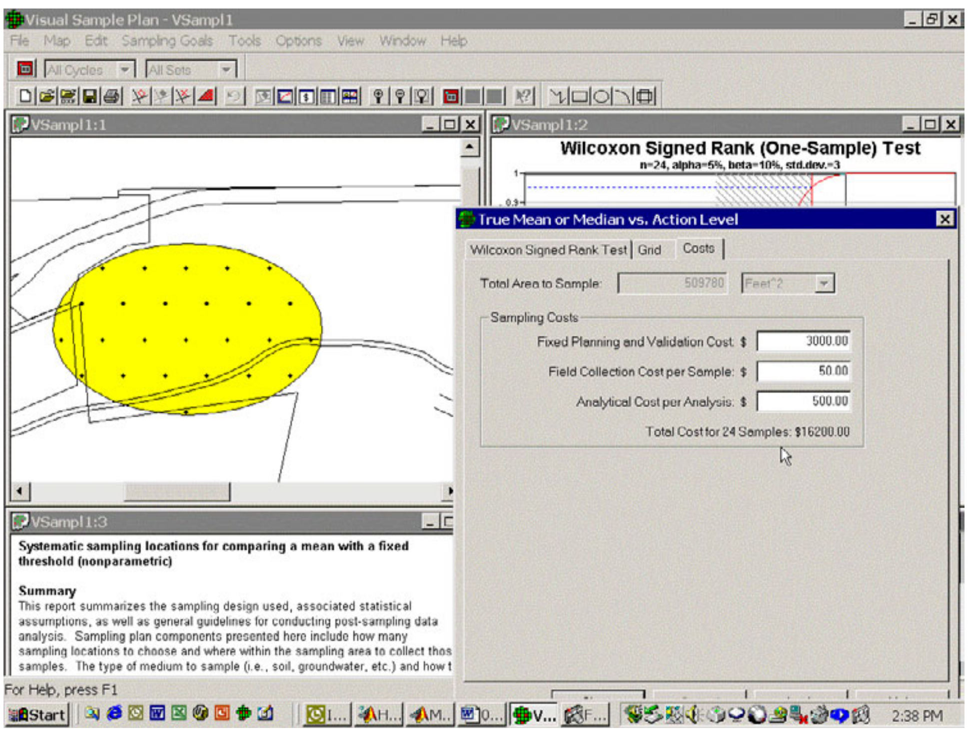


Figure 9. General tab for Wilcoxon Signed Rank Test with example costs

Goal: To learn how to export the coordinates of a sampling design to a .txt file.

- **Close** the **True Mean or Median vs. Action Level** dialog box.
- Enlarge the coordinates view, the lower-right window with X, Y locations, to fill the screen by clicking on its **Maximize** button or double clicking in its title bar.
- From the main menu select **Map→Sample Points,** and then highlight the **Export…** option.
- Click on the **Export…** option.
- At this point, your screen should look like Figure 10.
- A dialog box entitled **Save Sample Coordinates to a File** allows you to save the sample coordinates to a text file of your choice. These coordinates can be used in the field to determine the sampling locations.
- Press **Cancel** to close the dialog box.

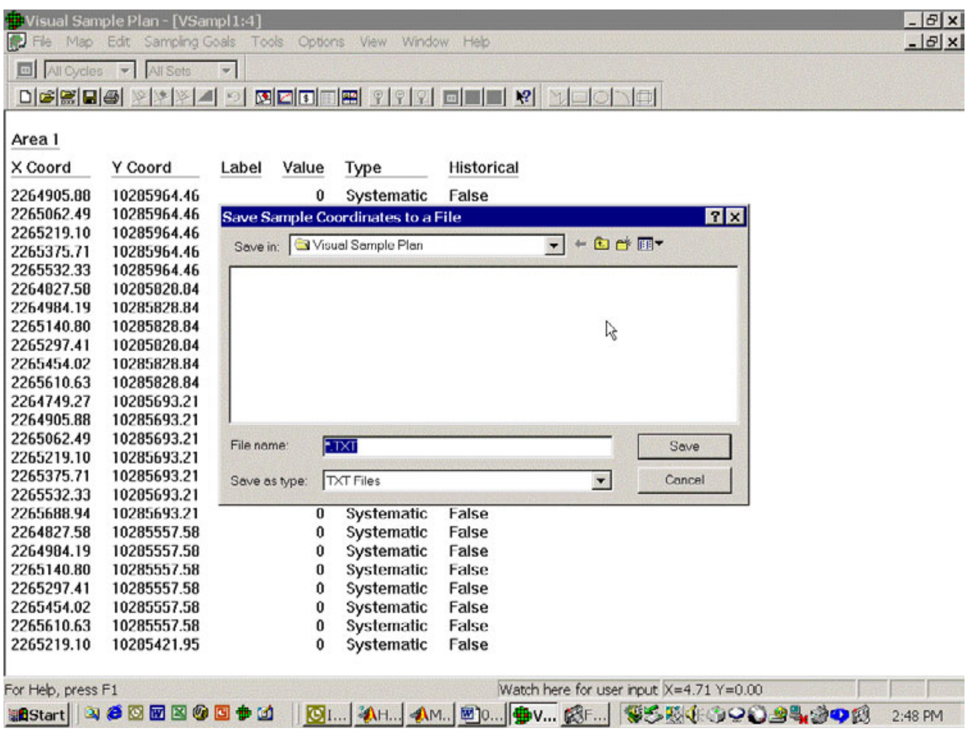
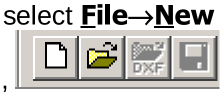


Figure 10. Coordinates view and Map→Sample Points→Export option

Visual Sample Plan Training Course
03 – VSP 2.0 Introductory Exercises

Presented by Bechtel Hanford, Inc.
Environmental Restoration Contractor

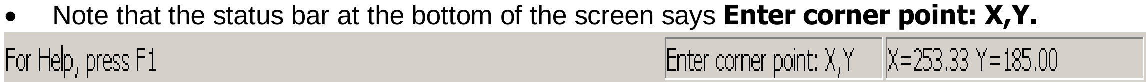
Goal: To open a new project, to draw a rectangle with specific dimensions, and then to create a sampling design suitable for a one sample t-test.

- Close the current project using **File→Close Project**. Respond **No** to the query **Save changes to Vsampl1?**
- From the main menu, select **File→New Project,**
- **Or** click the **New** icon, 

  ------▲
- Click the **Maximize** button (or double click title bar) to maximize initial window.
- Draw a one-acre field with overall dimensions of 165 feet x 264 feet = 43,560 square feet = 1 acre by doing the following:
  - Visual Sample Plan plots geometrical figures in an X-Y coordinate frame. For your one-acre field, it will be convenient to frame the field in your screen by changing the area portrayed on the screen. You can do this by setting the map extents for the figure you are about to draw. From the main menu select **Map→Set Map Extents.** A Map Extents dialog box will appear.
    - Using the dialog box, set map extents as follows:

      Minimum X: **-100** Minimum Y: **-100**
      Maximum X: **300** Maximum Y: **200**

      Use default **Units:** Feet
      Click **OK.** The dialog box now says **Sample Plan Phase 1**.
  - From the main menu select **Map→Draw Rectangle.**
  - Note that the status bar at the bottom of the screen says **Enter corner point: X,Y.**



▲

  - Using the keyboard, type in the coordinates of one corner of the one-acre field. Type (***be sure to use a comma*** between the numbers):
    0,0 <Enter>
  - The message on the status bar will change to **"Enter other corner: X, Y".** Now type:
    264, 165 <Enter>
    The one-acre field will appear in the color yellow.

- From the main menu select **View→Zoom Out**. Notice the cursor has changed to a magnifying glass  and the icon that looks like a magnifying glass with a minus sign in it is selected. The **Zoom Out** icon  is in the center.
- Shrink the one-acre field by clicking on the map three or four times.
- Deselect the **Zoom Out** icon by clicking on it.
- Now add a title: from the main menu select **Edit→Titles→Text…;** type in the title **One Acre Field**; click **OK.**
- From the main menu, select **Sampling Goals→Compare Average to Fixed Threshold Data not required to be normally distributed→Systematic grid sampling (Wilcoxon signed ranks test) ...**
- A **True Mean vs. Action Level** dialog box will appear. Be sure that it contains the following values:
- **Choose:** True Mean >= Action Level (Assume Site is Dirty).
- **False Rejection Rate (Alpha)** 5.0%
- **False Acceptance Rate (Beta)** 20.0%
- **Lower Bound of Gray Region** 9
- **Action Level** 10
- **Estimated Standard Deviation** 3
- Click **Apply.**

Note the bottom line in the dialog box in Figure 11: **Minimum Number of Samples in Survey Unit: 67.**

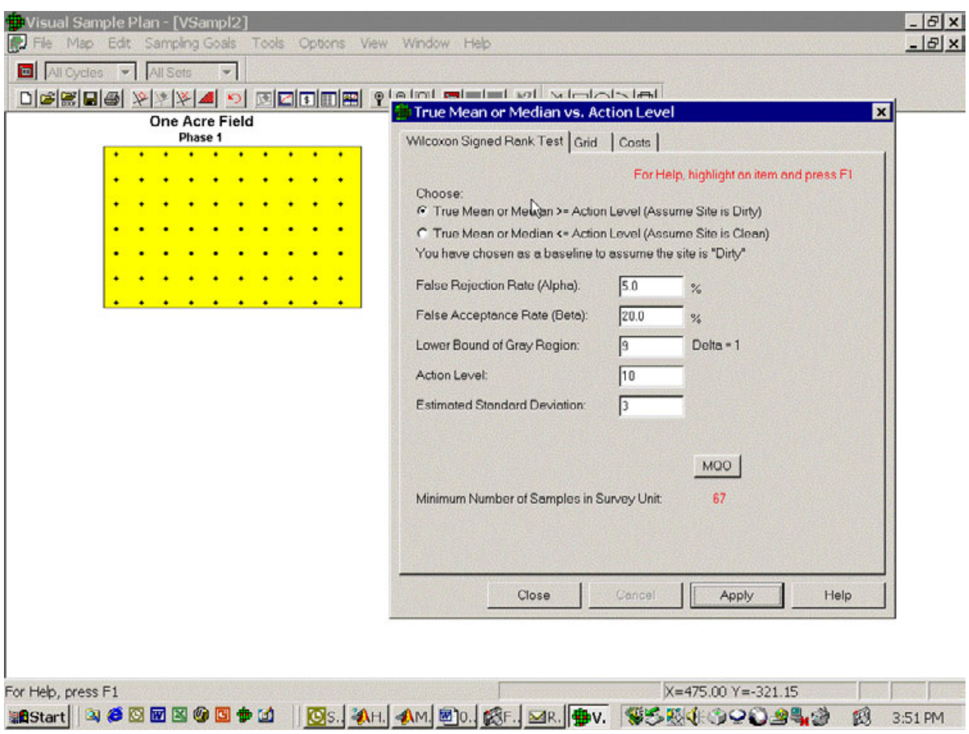


Figure 11. Minimum Number of Samples in Survey Unit: 67

Goal: To see the impact of altering the width of the gray region.

To make your screen look like Figure 12 below:

- Change the **Lower Bound of Gray Region** from **9 to 8.** Note that Delta, the width of the gray region, has increased from 1 to 2 units.
- Click **Apply.**
- Note **Minimum Number of Samples in Survey Unit: 18.**

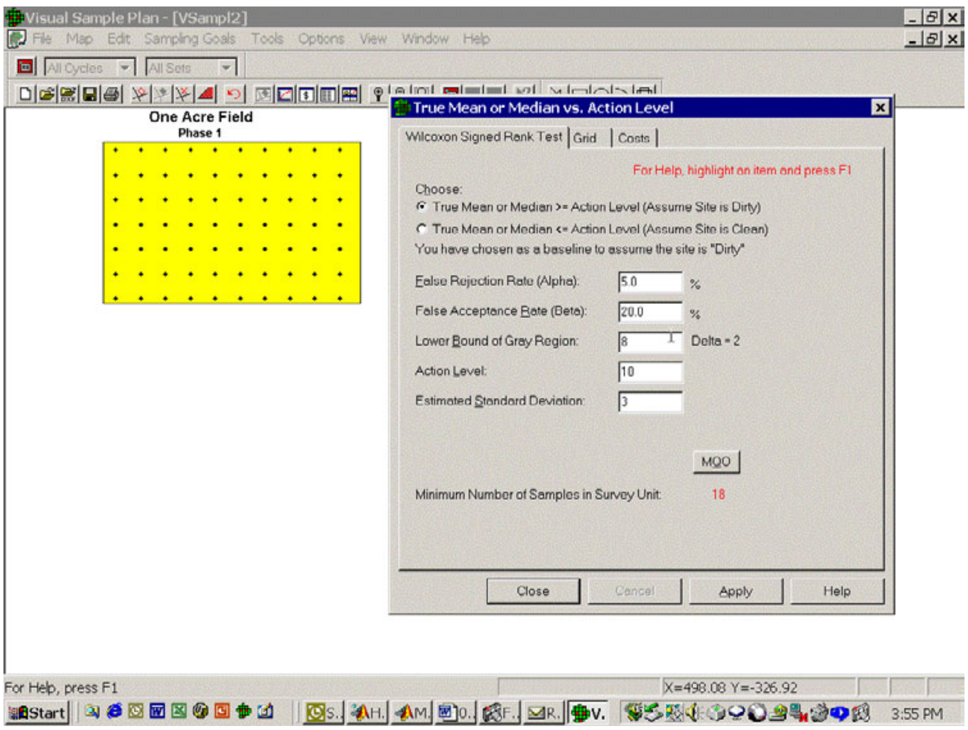


Figure 12. Dialog box with larger gray region

The dramatic decrease in sample size from 67 in the previous example to 18 in this example illustrates the important influence the width of the gray region has on sample size.

Close the dialog box.

Presented by Bechtel Hanford, Inc.
Environmental Restoration Contractor

Goal: To demonstrate the effect of changing the hot-spot size on the number of samples required.

To make your screen look like Figure 13 below:

- From the main menu select **Sampling Goals→Locating a Hot Spot→Systematic grid sampling→Minimize number of samples …** A dialog box will appear labeled **Find Hot Spot with Minimum Number of Samples.**
- Click the **Hot Spot** tab.
  - For **Shape (0.2 – 1.0)** enter **0.8.** (This command sets the assumed shape of the hot spot. A perfect circle has a shape of 1.0, a Shape = 0.8 ellipse has a minor-to-major axis ratio of 0.8.)
  - Enter **20.0** for the **Length of Semi-Major Axis** in **Feet**.
- Click the **Find Grid** tab within the dialog box and select **Triangular** as the **Grid Type.**
  - Enter **95%** for **Probability of Hit.**
  - Click **Apply.** Your screen will look similar to Figure 13.
- Click the **Close** button on the **Find Hot Spot with Minimum Number of Samples** dialog box.

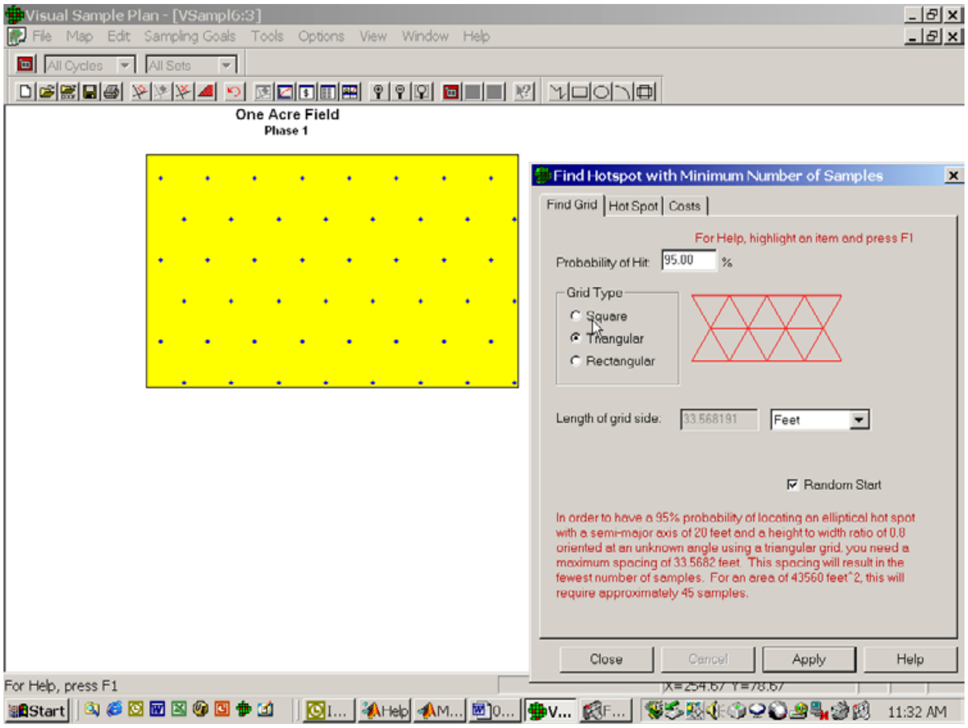
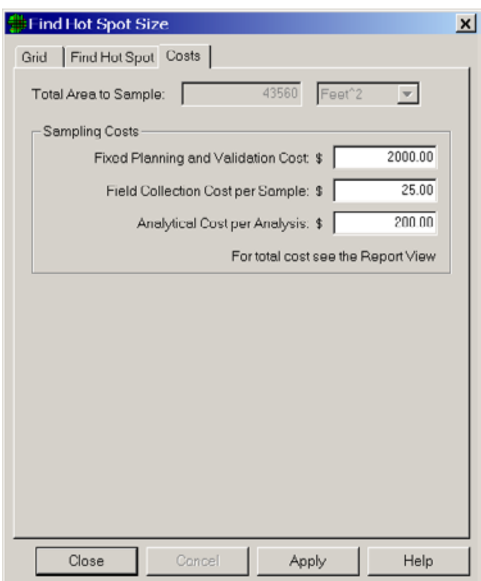


Figure 13. Hot-spot sampling design

Goal: To illustrate how to get sampling cost information with hot-spot sampling.

- From the main menu select **Window→Quad Window**.
- **Maximize** the report view in the lower, left corner of the screen.
- From the main menu select **Sampling Goals→Locating a Hot Spot→Systematic grid sampling→Minimize hot spot size …** A dialog box will appear labeled **Find Hot Spot Size.**
- Click the **Costs** tab and enter the following values:
  - Fixed Planning and Validation Cost **$2000.00**
  - Field Collection Cost per Sample **$25.00**
  - Analytical Cost per Analysis **$200.00**
- Click **Apply.** Click **Apply** several times. Note the **SUMMARY OF SAMPLING DESIGN** table; the entry "Number of samples on map" varies slightly, depending on the starting point of the sampling grid.
- Note the "Total cost of sampling" is listed as **$12125.00.**
- You can scroll the report window and see a large amount of information relating to the current sampling design.
- Click the **Close** button on the **Find Hot Spot Size** dialog box.



SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Detect the presence of a hot spot with a given size and shape
Type of Sampling Design	Hot spot
Sample Placement (Location) in the Field	Systematic (Hot Spot) with a random start location
Formula for calculating number of sampling locations	Singer and Wickman algorithm
Calculated total number of samples	45
Number of samples on map [a]	40
Number of selected sample areas [b]	1
Sampling area [c]	43560.00 ft^2
Size of grid / Area of grid cell [d]	33.5682 ft / 975.858 ft^2
Grid pattern	Triangle
Total cost of sampling [e]	$12125.00

Figure 14. Hot-spot sampling example

Goal: To develop a sampling plan with a limited budget.

To make your screen look like Figure 15 below:

- From the main menu select **Sampling Goals→Locating a Hot Spot→Systematic grid sampling→Predetermined fixed cost …**
- From the **Find Cost** tab, enter the **Total Cost** as **$10000.00**
- Verify that the **Grid Type** is still **Triangular**.
- The inputs under the **Hot Spot** and **Costs** tabs should default to the previous values. For example, the **Length of Semi-Major Axis** should still be 20.0 ft.
- Click **Apply.** Note (in the enlarged report window in the lower left-hand window of the **Quad Window**) that by fixing the total survey cost and the sample cost, we have reduced the number of samples to 35. However, the probability of hitting a hot spot has been reduced to about 80%.

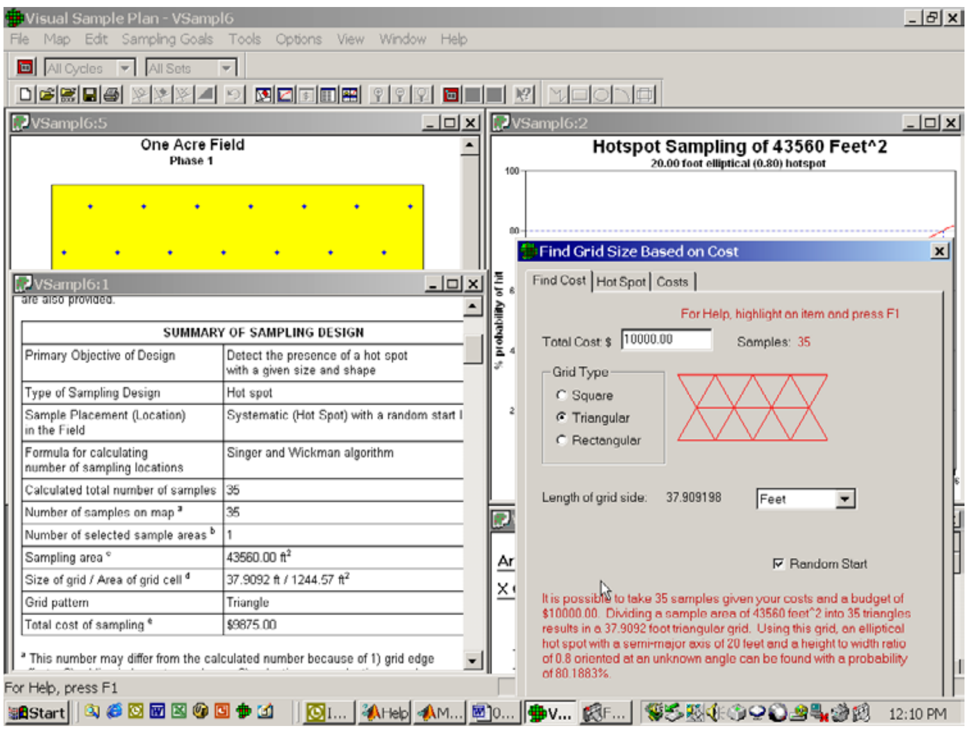


Figure 15. Hot-spot sampling option, "Predetermined Fixed Cost"

Although the sampling cost has been reduced to within the goal of $10,000, the 80% chance of detecting a hot spot may not meet our objectives. The next exercise will illustrate how to use the graph to compare sampling cost to the probability of hot-spot detection.

Goal: To see the graph view with probability of finding a hot spot vs. sampling cost.

- Enlarge the graph to fill the screen by clicking on its **Maximize** button. There is no need to close the dialog box at this point, but you may want to drag it off to one side.
  - From the main menu select **Options→Graph→Display Cost.** Note that the X-axis variable has changed from "Number of Samples using a triangular grid" to "Sampling Cost using a triangular grid". Also note that by moving the black line along the red curve, one can see the sampling cost for a probability of hit in the status box near the lower-right corner of the screen.

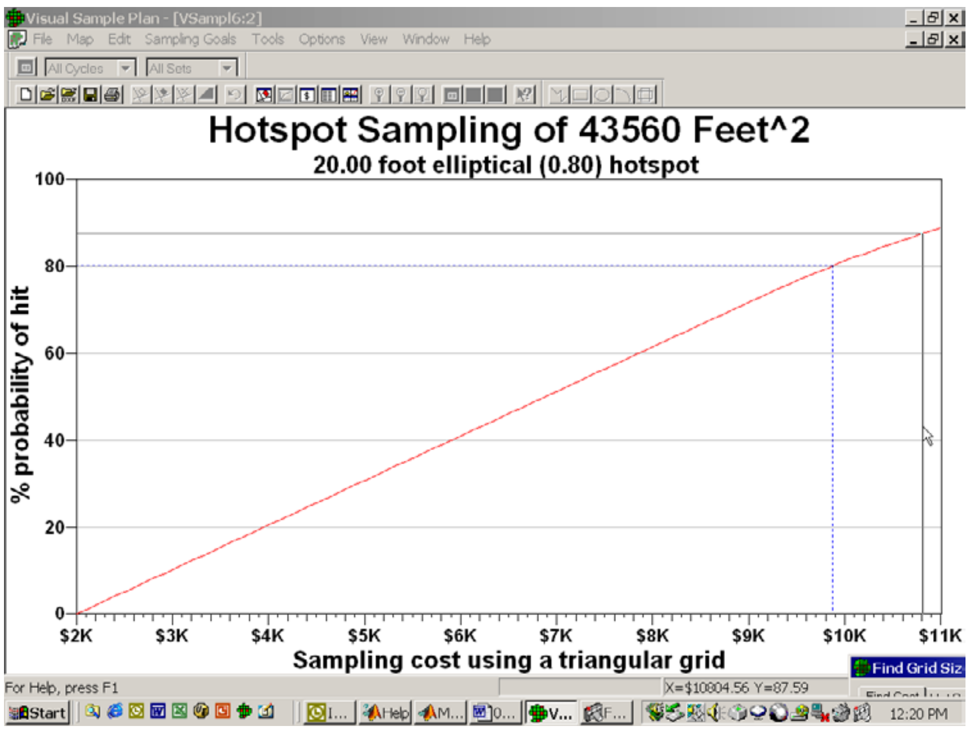


Figure 16. Hot-spot sampling example showing sampling cost on X-axis

In Figure 16, the status box shows that the cursor is at a sampling cost of $10,804.56 vs. a probability of hot-spot detection of 87.59%. Move the cursor and note the changes in the status box.

The dotted blue line goes vertically from the sampling cost that came within our budget of $10,000 ($9,875 actually selected by VSP) and then horizontally to the corresponding hot-spot detection probability of 80%.

For more information on the hot-spot problem, see *Statistical Methods for Environmental Pollution Monitoring* (Gilbert 1987, pp. 119-131).

Goal: To find the largest circular area that could be missed by a given systematic grid sampling design.

To make your screen look like Figure 17 below:

- Click the **Close** button in the **Find Grid Size Based on Cost** dialog box.
- Select the map view using **View→Map**.
- From the main menu select **Tools→Largest Unsampled Spot→Find…**
- Enter an **Accuracy** of **1.0** feet.
- Leave the two check boxes at their default values.
- Click **OK**. A dialog box will appear indicating that the largest potential circular area has a radius of about 22 Feet and an area of about 1,500 $Feet^2$. These values will vary depending on the current grid.
- Click **OK** to close the box giving the Radius and Area information.

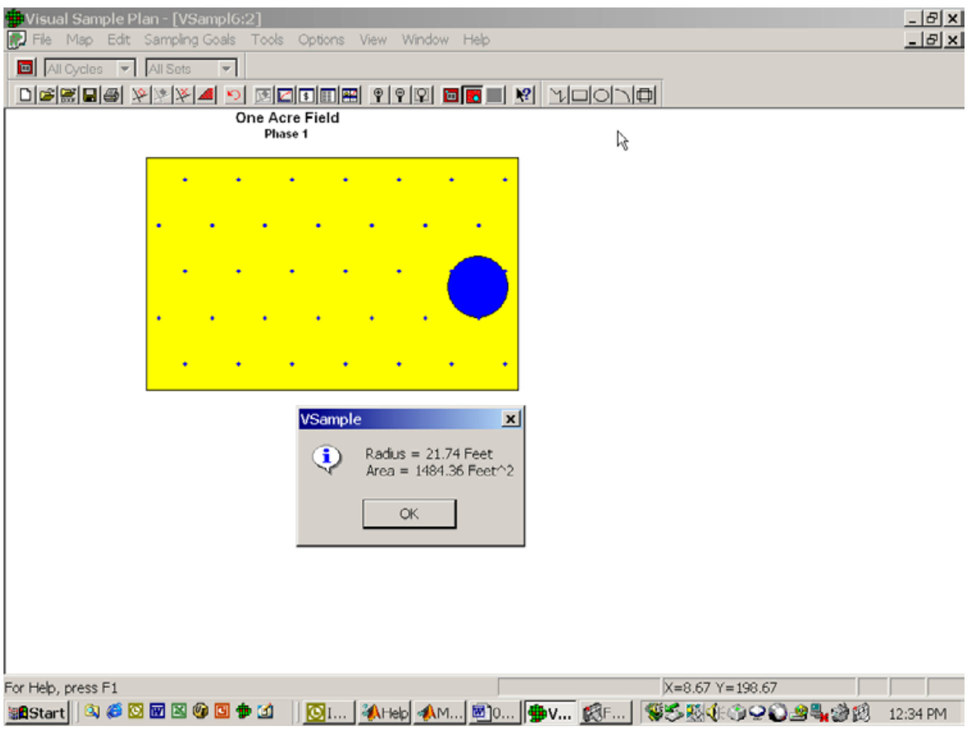


Figure 17. Largest potential unsampled circular area

The colored circle is the largest circle of contaminant that could be missed with the current sampling design. Note that no probability is associated with this tool. It simply finds the largest circle that fits between the current sampling locations.

Presented by Bechtel Hanford, Inc.
Environmental Restoration Contractor

Goal: To stratify a one-acre field for more efficient sampling.

To make your screen look like Figure 18 below:

- Close the project using **File→Close Project.** Respond **No** to the query **Save changes to VSampl2?** (Visual Sample Plan automatically numbers your projects as you open them. Thus you may have a VSampl3 or more at this point.)
- (Note: Visual Sample Plan supports a stratified random sampling design for determining the global average for a number of strata. However, our goal here is to divide a site into a number of smaller areas, each of which will be individually examined. The following example illustrates this subdivision approach.)
- From the main menu select **File→New Project.**
- Click the **Maximize** button (or double click title bar) to maximize initial window.
- From the main menu select **Map→Set Map Extents.** Set map extents as follows:
  Minimum X: **-100** Minimum Y: **-100**
  Maximum X: **300** Maximum Y: **200**
  Use default **Units**: Feet
  Click **OK**
- Divide the one-acre field into two subsections as follows:
  - From the main menu select **Map→Draw Rectangle**.
  - Using the keyboard, enter the coordinates of opposite corners of the first rectangular subsection (note that these entered coordinates appear in the status bar at the bottom of the screen): First type (note commas):
    0,0 <enter>
    then
    75,165 <enter>
  - Set the color of this subsection to dark blue using main menu option **Edit→Sample Areas→Change Colors.**
  - **Deselect** this subsection by left-clicking on it.
  - Start the second stratum by selecting **Map→Draw Rectangle.**
  - Using the keyboard, enter the coordinates of opposite corners of the second rectangular subsection (note that these entered coordinates appear in the bar at the bottom of the screen): First type (note commas):
    75,0 <enter>
    then
    264,165 <enter>
  - Set the color of this subsection to light green using main menu option **Edit→Sample Areas→Change Colors.**
  - Add a title: From the main menu select **Edit→Titles→Text …**; type in **One Acre Field** for the **Title** and **Two Strata** for the **SubTitle**. Click **OK.**
  - You should now have a sampling area divided into two subsections. Each subsection, strata, can now be treated separately.

  - The next sampling design will only apply to the right subsection in green. (Be sure the left subsection in blue is *not* selected by left-clicking on it if necessary.)

- From the main menu select **Sampling Goals→Compare Average to Fixed Threshold→Can assume data will be normally distributed→Systematic grid sampling …**
- A **True Mean vs. Action Level** dialog box will appear.
- Select the **Grid** tab and select **Triangular** grid type
- Select the **One-Sample t-test** tab. Be sure that it contains the following values:
  - **Choose:** True Mean >= Action Level (Assume Site is Dirty).
  - **False Rejection Rate (Alpha)** 5.0%
  - **False Acceptance Rate (Beta)** 10.0%
  - **Lower Bound of Gray Region** 5
  - **Action Level** 6
  - **Estimated Standard Deviation** 1.2
- Click **Apply.** This should leave a sparse triangular sampling grid with 14 points in the right-hand subsection.

- Left-click within the right-hand, green subsection to deselect it (remove color).
- Left-click within the left-hand, blue subsection to select it. The selected sampling area should now be colored blue.
- Be sure that the variables are set to the following values, noting that the standard deviation is now **2.4,** i.e., twice as large as above:
  - **False Rejection Rate (Alpha)** 5.0%
  - **False Acceptance Rate (Beta)** 20.0%
  - **Lower Bound of the Gray Region** 5
  - **Action Level** 6
  - **Estimated Standard Deviation** **2.4**
  - Click **Apply.** This should create a denser triangular sampling grid in the blue subsection.
- Left-click in the right-hand subsection to display it as well. Your screen should now closely resemble Figure 18.

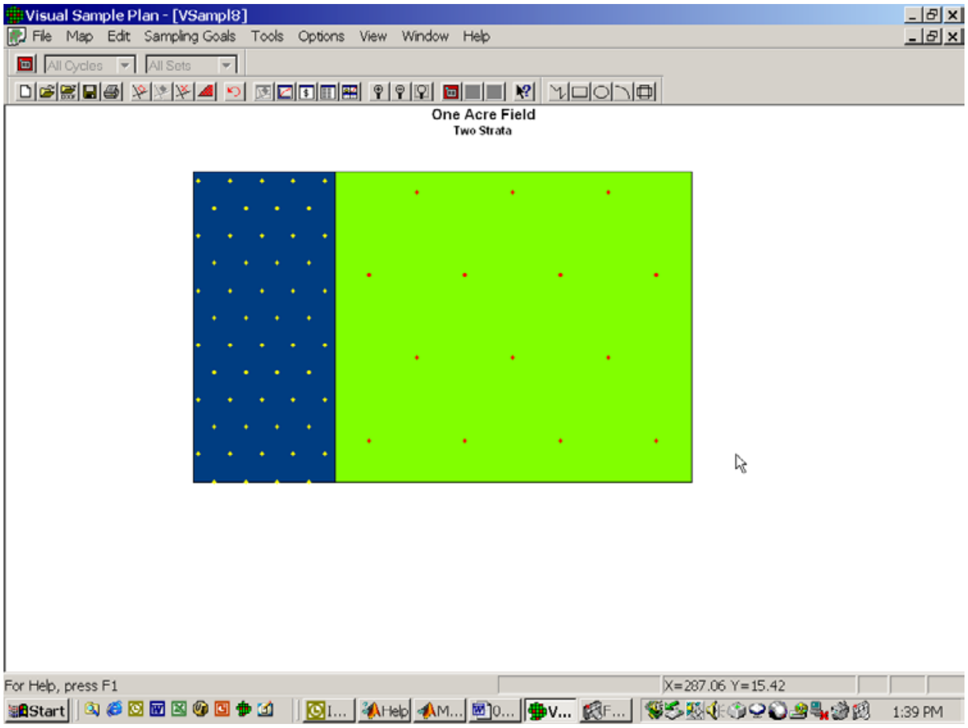


Figure 18. Example of subdivided sampling area

- Close the project using **File→Close Project.** Respond **No** to the query **Save changes to Vsampl3?**

Presented by Bechtel Hanford, Inc.
Environmental Restoration Contractor

Goal: To create a map of room surfaces (MARSSIM Room) that can be used for developing a sampling plan suitable for the Wilcoxon Rank Sum Test.

To make your screen look like Figure 19 below:

- From the main menu select **File→New Project.**
- Click the **Maximize** button (or double click title bar) to maximize initial window.
- From the main menu select **Map→Draw MARSSIM Room.**
  - Using the keyboard, type in the following dimensions (use **x** this time instead of comma): **60x50x40 <enter>**
  - A dialog box will appear with the question **Include ceiling?** Click **Yes.**
  - Select the **Zoom Out** icon. Click twice on the MARSSIM room to reduce its size. Deselect the **Zoom Out** icon.
- Remove the titles: From the main menu select **Edit→Titles→Text ...**; delete the existing title and sub-title; click **OK.**
- From the main menu select **Sampling Goals→Compare Average to Reference Average→Data not required to be normally distributed→Systematic grid sampling (MARSSIM WRS test) ...**
  - A dialog box will appear labeled **True Mean or Median vs. Background Level**.
  - Select the tab **Grid** tab and choose **Square** as the grid type.
  - Select the tab **MARSSIM WRS Test.** Be sure the list of parameters contains the following values:
    - **Choose:** True Mean >= Action Level (Assume Site is Dirty).
    - **False Acceptance Rate (Alpha)** 5.0%
    - **False Rejection Rate (Beta)** 5.0%
    - **Lower Bound of Gray Region** 3
    - **Specified Diff. of True Means or Medians [Action Level]** 5
    - **Estimated Standard Deviation** 2
  - Click **Apply.** This should leave a sampling grid with 33 sampling points.

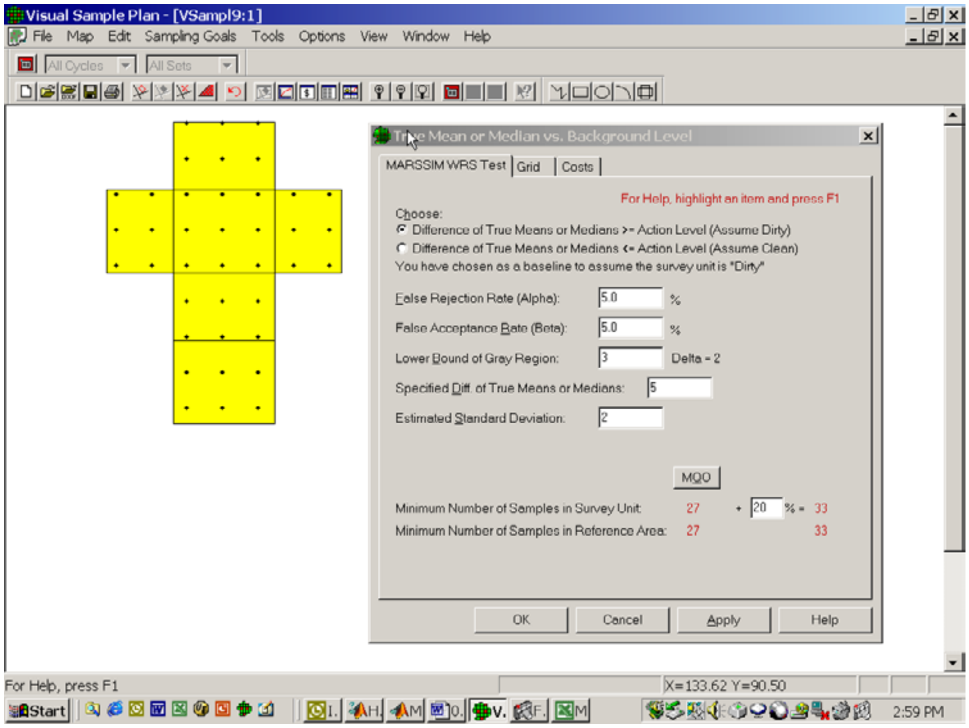


Figure 19. MARSSIM WRS Test dialog and room with square grid

Close VSP using **File→Exit**.
Click **No** to the query **Save Changes to Vsampl4**?

Goal: To gain more information about VSP.

For more specific information on VSP 2.0, see the *Visual Sample Plan Version 2.0 User's Guide* (Hassig, Wilson, Gilbert, Carlson, O'Brien, Pulsipher, McKinstry, and Bates 2002). It is available in PDF format at http://dqo.pnl.gov/VSP.

For background on the DQO and DQA process, see *Guidance for the Data Quality Objectives Process* (EPA 2000a) and *Guidance for Data Quality Assessment* (EPA 2000b). These documents are available in PDF format at http://www.epa.gov/quality1/qa_docs.html.

VSP 1.0 uses the *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)* (EPA 1997) document as the basis for MARSSIM calculations. Newer versions of MARSSIM in PDF format are available at http://www.epa.gov/radiation/marssim/obtain.htm.

Presented by Bechtel Hanford, Inc.
Environmental Restoration Contractor

References

Hassig, NL, JE Wilson, RO Gilbert, DK Carlson, RF O'Brien, BA Pulsipher, CA McKinstry, DJ Bates. 2002. *Visual Sample Plan Version 2.0 User's Guide*. PNNL-14002, Pacific Northwest National Laboratory, Richland, Washington.

EPA. 1997. *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*. EPA 402-R-97-016, NUREG-1575, U.S. Environmental Protection Agency, Washington, D.C.

EPA. 2000a. *Guidance for the Data Quality Objectives Process – QA/G-4*. EPA/600/R-96/055, Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C.

EPA. 2000b. *Guidance for Data Quality Assessment – Practical Methods for Data Analysis – QA/G-9, QA00 Update*. EPA/600/R-96/084, Office of Environmental Information, U.S. Environmental Protection Agency, Washington, D.C.

Gilbert, R. O. 1987. *Statistical Methods for Environmental Pollution Monitoring*. John Wiley & Sons, Inc., New York.

Example 1: The Effects of α, β, and σ on Sample Size

Objective:

(1) To familiarize attendees with creating a statistical sampling design using simple random sampling techniques.

(2) To demonstrate how changes in α, β, and σ affect the number of samples required. (Note that while this exercise uses parametric statistics, the effects of α, β, and σ are similar when using non-parametric statistics.)

Background Information:

The mission of the site under investigation was to produce components for nuclear weapons from materials such as plutonium, uranium, and beryllium. The primary processing plant is located near a one-acre field that was part of a buffer area surrounding the plant. In 1946, a fire in the primary processing plant released 30 millicuries of Pu-239. The one-acre field is now being considered for release for industrial land use, but it has never been characterized and is presumed to be contaminated (i.e. Ho = site is contaminated). The site manager has asked you to calculate the number of samples required to characterize the site for Pu-239. The regulators have decided that if the average Pu-239 level on the site is greater than 8 pCi/g, then the site will need to be remediated. They prefer limiting the chance that a contaminated site will not be remediated to 1%, however they have agreed, if necessary, to accept a 5% chance that a contaminated site will not be remediated (that is α can be set at 1% or 5%). The site manager has decided that he can live with a 15% to 20% probability of remediating a clean site (that is, β can be set at 15% or 20%). Three historical data sets from similar sites are available for estimating the value of the standard deviation of plutonium soil concentrations. One estimates the standard deviation to be 1.0 pCi/g, another at 2.0 pCi/g, and the third at 3.0 pCi/g. The Lower Bound of the Gray Region is set to 7 pCi/g (so the Width of Gray Area or Region is 1.0 pCi/g).

Exercise:

Step 1:

- Close any open windows by clicking "**x**".
- From the main menu, select **File** → **Open Project** → **OneAcreField.simple.vsp** → **Open**.
- Maximize the **OneAcreField.simple.vsp** box to fill the screen.

Step 2:

- From the main menu select **Sampling Goals** → **Compare Average to Fixed Threshold** → **Can assume data will be normally distributed** → **Simple random sampling…** → **True Mean vs. Action Level**. Use the tab **One-Sample t-Test**. Make sure that **True Mean >= Action Level (Assume Site is Dirty)** is selected as the null hypothesis.

E1-1 Using the information provided above, complete the following table, where n is the number of samples required or **Minimum Number of Samples in Survey Unit**. (Hint: to save time, click **Apply** after each parameter change.)

	α =		α =	
	β =	β =	β =	β =

σ =	n=	n=	n=	n=
σ =	n=	n=	n=	n=
σ =	n=	n=	n=	n=

Step 3:

In the space below, briefly answer the following questions:

E1-2 When n = 31 and σ = 2, the decision-maker is willing to walk away from a dirty site ____% of the time. What type of error is this?

E1-3 When n = 31 and σ = 2, the decision-maker is willing to clean up a clean site _____% of the time. What type of error is this?

E1-4 Which of the three parameters (α, β, or σ) has the biggest impact on the number of samples required?

E1-5 If σ is incorrectly underestimated, what is the impact?

Step 4:

- Close any open windows. Respond **No** to the query **Save changes to OneAcreField.simple.vsp?**

Example 2: Judgmental Sampling versus Simple Random Sampling

Objective:

(1) To illustrate the advantages and disadvantages of sampling plans using a fixed sample size for random sampling and for judgmental sampling.

Background Information:

A warehousing facility is being sold, and the prospective buyer is conducting an investigation to characterize existing environmental conditions and associated potential liability. One feature being assessed is a 5,000 ft^2 fenced area where drums of uranium oxide were stored temporarily during shipment. Some drums were apparently damaged during handling, leading to release of small amounts of material assumed to be yellowcake which has stained the soil yellow. However, there is no information on what contaminants the stains may contain. Eight yellowish stains have been identified, and a typical stain is about 1 ft^2.

Exercise:

Step 1:

- From the main menu, select **File → Open Project → drums.vsp → Open**. If necessary, maximize the **drums.vsp** window.

Step 2:

Select the eight yellow stains as follows:

- From the tool bar select the **Zoom In** button (Ensure that the **Main Toolbar** is activated). Using the magnifying glass cursor, click on the fenced area 5 times so that it to nearly fills the screen. Use the horizontal and vertical scrolls bars to center the fenced area. Deselect the **Zoom In** button by clicking on it.

- From the main menu select **Edit → Sample Areas → Select/Deselect Sample Areas**.

- A **Select / Deselect Areas** text box will appear. Click the **Deselect All** button to deselect all sample areas and then click **OK** to close the text box. The background gray color of all previously selected areas will disappear, leaving only the outline of the area.

- Now repeat the step of selecting **Edit → Sample Areas → Select/Deselect Sample Areas**. This time Highlight **Area 118** through **Area 125**; click **OK**. The eight ellipses will be filled in with a yellow color to form the eight stained areas

Step 3:

- Use the **Zoom In** button from the tool bar to fill the screen with the left half of the fenced area where the drums are located. Then deselect the **Zoom In** button. Use the horizontal and vertical scroll bars to move the section with the drums so that they are all in view.

- From the main menu select, **Sampling Goals → Non-statistical sampling approach→Judgement (authoritative) Sampling.** Note that the cursor has now changed to a crosshair. Click on each stained area to select one sample point within it.

Step 4:

If these eight samples constitute your sampling plan, answer the following questions (without using the computer):

E2-1 For the purpose of characterizing the eight stained areas, is this judgmental design appropriate? Explain why or why not.

E2-2 For the purpose of selling the site (and transferring any related risk), can you extrapolate data from the judgmental sampling design throughout the fenced area? Explain why or why not.

E2-3 Using the data from the judgmental sampling design, can conclusions made about this site be extended to analogous sites?

E2-4 For this sampling design, were Type I and Type II errors identified and controlled?

Step 5:

- Deselect the eight stained areas: from the main menu select **Edit** → **Sample Areas** → **Select/Deselect Sample Areas**. Deselect **Area 118** through **Area 125** by clicking the button **Deselect All**; then click **OK**. The eight yellow ellipses will disappear.

- Then select the entire fenced area: from the main menu select **Edit** → **Sample Areas** → **Select/Deselect Sample Areas**. Highlight **Area 1**; then click **OK**. The fenced area will be highlighted in gray.

- From the main menu, select **Sampling Goals** → **Non-statistical sampling approach** → **Predetermined number of samples** → **Simple random sampling….** Enter **8** in the **Number of Samples** box under the **Samples** tab. This option randomly places your prespecified or predetermined number of samples within the fenced area. You may want to move the **Simple Random Sampling** text box aside so that you can see the sample points more clearly. Click **Apply** and notice where the sampling points are located on the map. (Click **Apply** several times to see different random patterns of sampling points.)

- Click the **Close** button to close the text box.

Step 6:

Answer the following questions:

E2-5 How are the judgmental and predetermined sampling designs different? (Hint: Think beyond the physical layout of the sample points.)

E2-6 For the predetermined sampling design, were Type I and Type II errors controlled?

E2-7 Can these conclusions be extended to analogous sites?

Step 7:

- Now select **Sampling Goals →Compare Average to Fixed Threshold → Data not required to be normally distributed → Simple random sampling (Wilcoxon signed ranks test)… → True Mean or Median vs. Action Level.** Assume that null hypothesis is that the site is dirty. Under the **Wilcoxon Signed Rank Test** tab, begin by setting α = 5% and β= 20%. Assume that the **Action Level** is **10** mg/kg, the **Width of Gray Region (Delta)** is **2.5** mg/kg, and the **Estimated Standard Deviation** is **6** mg/kg.

E2-8 What combination of α and β are needed to get 8 samples? (Note: There may be more than one correct answer.)

Step 8:

E2-9 Are these tolerable levels for decision errors? That is, will decision-makers find these acceptable levels for walking away from a dirty site or cleaning up a clean site? Explain why or why not.

- Now select appropriate values for α and β, insert them in the Text box, and click **Apply** to get the required sample size.

E2-10 How does this compare with the fixed sample size of 8?

Step 9:

- Close any open windows. Respond **No** to the query **Save changes to drums.vsp?**

Example 3: Parametric versus Non-Parametric Statistics

Objective:

(1) To compare the impact of parametric versus non-parametric statistics in the number of samples required to characterize a site.

Background Information:

From 1915 to 1927, a New Jersey clock and instrument company manufactured radiant dials, using radium paint to hand-paint the dials. The small factory was severely contaminated (Ho: site is contaminated), leading to a series of demolition and decontamination efforts over several decades as radiation protection standards changed. At present, the site is now a vacant lot where a pile of rubble mixed with sand and gravel still remains. Responsibility for the site has passed to an EPA Region 2 site manager who must decide whether to send the pile of rubble, sand, and gravel to a public landfill or to a licensed radioactive waste disposal facility in Utah. The decision will be based on the residual radium-226 concentration above the site background. To dispose of the rubble at a (much less expensive) public landfill, the EPA site manager must demonstrate that the residual radium-226 concentration in the mixed rubble is no more than 5 pCi/g. The regulators want to be very sure that they do not release a contaminated site and have set α at 1%. The site manager is very anxious to avoid public controversy and err on the side of caution (sending the rubble to the more expensive Utah facility), so she is willing to send clean rubble to Utah 25% of the time (i.e., when the true mean is less than background). The width of the gray area is 1.6 pCi/g. Historical data suggests that the standard deviation for the radium is 2.5 pCi/g. The rubble pile is an ellipse with an area of approximately 740 square feet.

Exercise:

Step 1:

- From the main menu, select **File → Open Project → DialFactory.vsp → Open**. If necessary, maximize the **DialFactory** box.

Step 2:

- From the main menu select **Sampling Goals → Compare Average to Fixed Threshold → Data not required to be normally distributed → Simple random sampling (MARSSIM sign test)…**.

E3-1 How many samples will be required for rubble pile? (Hint: in this example, VSP functions better if you begin by entering the standard deviation.)

- Click the **OK** button to close the text box.

Step 3:

- A statistician tells the cleanup contractor that there are two options analyzing the data: parametric and non-parametric tests. Parametric tests are more powerful because they are based distributional assumptions. These assumptions allow the statistician to assign probabilities to the occurrence of events. However, before parametric statistics can be used, the distributional assumptions must be verified. Non-parametric statistics are based on much less stringent assumptions but are not as powerful and require more samples.

E3-2 Statistical tests have determined that approximately 30-50 samples are needed verify the distributional assumptions of a parametric test. Given this, determine whether it is worthwhile (in this case) to verify the distributional assumptions of a parametric test. (Hint: from the main menu select **Sampling Goals → Compare Average to Fixed Threshold → Can assume data will be normally distributed → Simple random sampling**)

Step 4:

E3-3 A new Regional Administrator drastically cuts the Superfund budget, leading the site manager to decide suddenly that she wants a much smaller chance of erroneously sending clean rubble to Utah. So she lowers β to 15%. How many samples will now be needed with either the MARSSIM Sign Test or the One Sample t-Test?

Step 5:

- **Close** any open windows. Respond **No** to the query **Save changes to DialFactory.vsp?**

Example 4: One Sample Proportion Test

Objectives:

(1) To provide practice in choosing tolerable decision error limits for a One Sample Proportion Test.

(2) To illustrate that VSP can handle a situation where action levels are not in the typical concentration units.

Background Information:

Ten thousand (10,000) cans of uranium oxide (UO_2) powder are stored in a warehouse. Each can weighs approximately 10 kg. A can is termed "defective" if its weight is not within 100 g of the value in the records system. If 20% or more of the cans are found to be defective, then all the cans will require repackaging.

Your task is to advise the plant manager of the minimum number of drums that need to be inspected to estimate the percentage of defective cans in the population. The consequences of leaving defective cans in-place without repackaging has a very low risk. Repackaging however is very expensive, so she wants to guard against unnecessarily repackaging the cans.

Exercise:

Step 1:

From the main menu, select **File** → **Open Project** → **warehouse.vsp** → **Open**.

Step 2:

E4-1 What is the Null Hypothesis for this project?

E4-2 What is the Action Level for this project?

E4-3 Why is the Action Level for this project high?

E4-4 How is this Action Level different from those normally encountered in soil remediation or D&D projects?

E4-5 Given the Null Hypothesis: The proportion of defective cans in the warehouse is ≥ 20%; would you advise the plant manager to set the Type I error to 1%, 5%, 10% or 20%? Explain your answer.

E4-6 Given the Null Hypothesis: The proportion of defective cans in the warehouse is ≥ 20%; would you advise the plant manager to set a Type II error to 1%, 5%, 10% or 20%? Explain you answer.

From main menu, select **Sampling Goals →Compare Proportion to Fixed Threshold → Data not required to be normally distributed → Simple random sampling**.

E4-7 Given the Null Hypothesis: The proportion of defective cans in the warehouse is ≥ 20%; Type I error = 20%, Type II error = 1%, and the Width of the Gray Area (Region) is 0.1, what is the minimum number of drums that need to be inspected to check the Null Hypothesis.

Step 4:

Close any open windows. Respond **No** to the query **Save changes to warehouse.vsp?**

Example 5: Finding Hot Spots

Objective:

(1) To illustrate how to create sampling designs for locating hot spots.

Background Information:

Ten (10) one-acre sites in the 100 Area have been remediated down to 5 feet below grade. Before remediating further, the contractor wants to know if there are any remaining circular hot spots with a radius greater than 10 feet (Ho: Site is contaminated). The contractor decides to conduct an assessment of one of the one-acre plots.
He plans to send the samples he collects to a reputable laboratory for analysis. In addition to locating hot spots, the data will also be used to estimate the variance of the contaminant of concern at the site.

Exercise:

Step 1:

In the main menu, select **File** → **Open Project** → **100AreaField.simple.vsp**

Step 2:

From the main menu, select **Window** → **Quad Window**. The map you loaded should appear in the upper left-hand window. The other three windows are essentially blank at this point.

Step 3:

From the main menu, select **Sampling Goals** → **Locating a Hot Spot** → **Systematic grid sampling** → **Minimize hot spot size….** (Ensure that the **Random Start** box is checked in the box **Find Hot Spot Size** under the **Grid** tab.)

Step 4:

Click the **Find Hot Spot** tab, if is not already selected. Enter **95%** as the **Probability of Hit.** Click the **Grid** tab**.** Select **Triangular** as the **Grid Type.**

Click the **Costs** tab. The default value for **Fixed Planning and Validation Cost** is **$0.00**. The default value for **Field Collection Cost per Sample** is **$100.00** and the default value for **Analytical Cost per Analysis** is **$400.00**. Click **Apply**.

Review the information presented in the Report Window (the lower left-hand window).

Step 5:

Answer the following questions:

E5-1 How many samples will be required to have 95% confidence that a hot spot with a 10-ft radius or larger will be detected?

E5-2 What is the required spacing for the triangular grid?

E5-3 How much will it cost to have 95% confidence that a hot spot with 10-ft radius (or larger) will be detected? (Remember that the numbers are generated for a single one-acre site.)

E5-4 If none of the criteria (95% confidence and 10-ft radius of the hot spot) can be relaxed, what other options are open to the contractor?

E5-5 Change the size of the hot spot to have a radius of 20-ft. How does this affect the total cost if the confidence remains at 95% and a triangular grid is used?

E5-6 Change the shape of the hot spot by entering **0.6** in the **Shape (0.1 – 1.0)** box under the **Find Hot Spot** tab, keeping the other parameters the same. Set the **Length of the Semi-Major Axis** to ≈**10** ft by changing the **Length of grid size** (for triangular **Grid Type** under the **Grid** tab) to **14.0 ft**. How does it affect the total cost if the confidence remains at 95% and a triangular grid is used?

Step 6:

Close any open windows. Respond **No** to the query **Save changes to 100AreaField.simple.vsp?**

Example 6: What if I don't know anything about the site?

Objective:

To show how VSP can be used to create sampling designs even when there is no historical data about the site.

Background Information:

The site of a former low-level liquid waste evaporation pond in the Idaho National Environmental and Engineering Laboratory is being examined to determine the need for remediation. Process knowledge is available to determine the species and activities of radionuclides, which flowed into the pond, but no soil or sediment sampling has been performed as yet.

The action level has been set at 3.5 pCi/kg for the total gamma-emitting radionuclides, α and β are set at 5% and 20% respectively, and the lower bound of the gray area is 3 pCi/kg. However there is no historical information available with which to estimate the standard deviation.

Exercise:

Step 1:

- In the main menu, select **File → Open Project → evaporationpond.vsp → Open**. Then select **Sampling Goals → Compare Average to Fixed Threshold → Can assume data will be normally distributed → Simple random sampling…**.

Step 2:

E6-1 One relatively conservative approach **is to assume that the standard deviation is equal to the action level**. This approach can be used when there is no information whatsoever about the contaminant of concern. Using this assumption, find the number of samples needed for a parametric test.

Step 3:

- If the range (the maximum minus the minimum activity or concentration) can be estimated, and if the general shape of the population distribution is known, then several other approaches are possible. Assume that the likely minimum value is 0.5 pCi/kg and the likely maximum value is 6.5 pCi/kg.

- Please note that the relationship between range and standard deviation of a data set is discussed in *Some Theory Sampling, by William Edward Deming, Dover Publications, Inc, New York, 1950.*

E6-2 If the shape is **completely unknown or if a bimodal distribution (background plus contamination)** is expected, then estimate the standard deviation as the **range divided by 2.8.** Using this assumption, find the number of samples needed for a parametric test.

E6-3 If the population distribution is **uniform,** then estimate the standard deviation as the **range divided by 3.5.** Using this assumption, find the number of samples needed for a parametric test.

E6-4 If the standard deviation is **a right or left triangular (extremely positively or negatively skewed, respectively),** then estimate the standard deviation as the **range divided by 4.** Using this assumption, find the number of samples needed for a parametric test.

E6-5 If the population distribution is **pyramidal**, then estimate the standard deviation as the **range divided by 5.** Using this assumption, find the number of samples needed for a parametric test.

E6-6 If the population distribution is **bell-shaped**, then estimate the standard deviation as the **range divided by 6.** Using this assumption, find the number of samples needed for a parametric test.

E6-7 Is it worthwhile to investigate historical data to try to estimate the standard deviation?

Step 4:

Close any open windows. Respond **No** to the query **Save changes to evaporationpond.vsp?**

Example 1: The Effects of α, β, and σ on Sample Size

E1-1

	α = 1.0%		α = 5.0%	
	β = 15.0 %	β = 20.0 %	β = 15.0 %	β = 20.0 %
σ = 1	n = 15	n = 13	n = 9	n = 8
σ = 2	n = 48	n = 43	n = 31	n = 27
σ = 3	n = 105	n = 94	n = 67	n = 57

E1-2

When n = 31, σ = 2, and β = 15.0% the decision-maker is willing to walk away from a clean site **5%** of the time. This is an α or **Type I** error.

E1-3

When n = 31, σ = 2, and α = 5.0% the decision-maker is willing to clean up a clean site **15%** of the time. This is an β or **Type II** error.

E1-4

The **standard deviation** (σ) has the biggest impact on the number of samples required. Notice as you go down the columns, σ increases at a steady rate from 1 to 2 to 3. The number of samples required, however, does not follow a similar pattern. Instead, a unit change in σ causes an exponential change in the number of samples required.

E1-5

As illustrated in the example above, a small underestimation of σ can lead to a significant underestimation of the number of samples needed to meet the required levels for α and β (the tolerable limits on decision errors). If too few samples are taken and the required error tolerances are not met, the actual error tolerances will be unknown or uncontrolled. In other words, the uncertainty of the study being conducted will not be managed.

Example 2: Judgmental Sampling versus Simple Random Sampling

E2-1

The appropriateness of a sampling plan always depends on the desired use of the data. For the purpose of characterizing the eight stained areas in this example, taking eight judgmental samples (one from each stain) is appropriate. The reason that a judgmental sampling design is appropriate here is that the goal is simply to characterize the eight stains and not to make inferences beyond the specific sites (the eight stains) that were sampled.

E2-2

For the purpose of selling the site, the surface soil within the entire fenced area needs to be characterized. So the question becomes, can the eight judgmental samples be used to characterize the soil within fenced area as a whole? The answer is No. The reasoning is as follows: Each of the eight stained areas had a 100% chance of being sampled and the remaining surface soil with the fenced area had a 0% chance of being sampled. This being the case, the population from which the eight samples were drawn is confined to the eight stains themselves. The rest of the soil within the fenced area is outside the population for this sampling scheme. From a statistical standpoint, it is inappropriate to make inferences to areas outside of the population being sampled.

EPA QA/G-5 (page 20) provides helpful guidance regarding judgmental sampling and inferences:

> "Judgmental sampling does not allow the level of confidence (uncertainty) of the investigation to be accurately quantified. In addition, *judgmental sampling limits the inferences made to the units actually analyzed*[emphasis added], and extrapolation from those units to the overall population from which the units were collected is subject to unknown selection bias."

E2-3

No, for reasons similar to those stated above. The analogous site is outside the population being sampled. Therefore, it would be inappropriate to make inferences to analogous sites, because all samples within the analogous site have 0% chance of being sampled with the current judgmental sampling design.

E2-4

When conducting judgmental sampling, there is no way to identify or control Type I and Type II errors. In other words, the probability of walking away from a dirty site and the probability of cleaning up a clean site are unknown and uncontrolled.

E2-5

As discussed above, judgmental sampling does not allow inferences beyond the specific samples that were taken. In addition, with judgmental sampling there is no control of Type I and Type II errors. The case where a predetermined number of samples are randomly placed within an area is somewhat more complicated. Because the samples are randomly placed, it is correct to call the design a statistical (rather than judgmental) sampling design. However, taking a predetermined number of samples does not allow the Type I and Type II error rates to be known and controlled a priori.

E2-6

No, the Type I and Type II errors were not controlled or managed.

E2-7

No, the conclusions drawn from the eight predetermined samples cannot be extended to analogous sites. The reasoning is similar to that in answer **E2-2**. The analogous site is not part of the population being sampled because all possible samples within the analogous site will not have a chance of being sampled.

E2-8

If α = 30.0% and β = 30.0%, the number of samples required will equal 8. There may be other combinations of α and β that will produce the same result.

E2-9

In this example, the buyer is most heavily impacted by a high α level. If α = 30%, the there is a 30% chance that the buyer will believe that the fenced area is clean when it is actually contaminated. Said another way, there is a 30% chance that the seller will walk away from a dirty site. It is doubtful that the buyer would find this acceptable. The seller is most heavily impacted a high β level. If β = 30%, then there is a 30% chance that seller will believe that the fenced area is contaminated when it isn't. Said another way, there is a 30% chance that seller will be required clean up a clean site. Again it is doubtful that the seller would likely find this acceptable. It should be noted, however, that the acceptability of tolerable decision error rates is dependent upon the particular situation.

E2-10

If smaller values for α or β are inserted in the text box, the resultant number of samples will be larger than 8. If larger values for α or β are inserted in the text box, the resultant number of samples will be smaller than 8.

Example 3: Parametric versus Non-parametric Statistics

E3-1

The number of samples required is 48. (where the numerical values of α = 1.0%, β = 25.0%, δ = 1.6, σ = 2.5, and Action Level = 5

E3-2

If a parametric test were performed, it would require 25 samples and the distributional assumptions would need to be tested. If the non-parametric test were performed it would require 48 samples and there would be no need to test the distributional assumptions. Parametric tests have more power to reject the null hypothesis on the basis of fewer samples. So in this case, it would be worthwhile to take 5 additional samples (resulting in 25 + 5 = 30 samples) in order to test the distributional assumptions and possibly be able to use the parametric test. If the distributional assumptions are verified, then the 30 samples already taken are sufficient. If the distributional assumptions are not verified, then the non-parametric test will need to be used and 18 additional samples will need to be taken for a total of 48 samples.

E3-3

The number of samples required with the MARSSIM Sign Test is 60. The number of samples required with One Student t-Test is 31.

Example 4: One Sample Proportion Test

E4-1

The Null Hypothesis for the project is Ho: $P \geq 0.2$ (i.e. the proportion of defective cans in the warehouse is ≥ 20%).

E4-2

The Action Level for this project is 20% of the can inventory are defective.

E4-3

The Action Level illustrated in this example is high because the consequences of leaving the defective cans in place without repackaging has very low risk as the warehouse is not actively used and the consequences of (unnecessary) repackaging is very expensive.

E4-4

The Action Level for this project is a proportion, while the Action Levels normally encountered in soil remediation or D&D projects are concentrations or activity levels.

E4-5

The plant manager would be advised to set the Type I error, α to be 20%, because the consequences of leaving the defective cans in place without repackaging has very low risk.

E4-6

The plant manager would be advised to set the Type II error, β to be as low as possible = 1%, because repackaging is expensive and unnecessary repackaging is wastefully expensive.

E4-7

The minimum number of drums to be inspected under the given conditions is 108.

Example 5: Finding Hot Spots

E5-1

When the **Random Start** option on the **Grid** tab is checked, the number of samples suggested by VSP can vary between 126 and 135 depending on how the **Triangular Grid Type** fits into the one-acre site.

If the **Random Start** option in the **Grid** tab is turned off, about 126 samples will be required to have 95% confidence that a hot spot with a 10-foot radius or larger will be detected.

The smallest circular hot spot that can be detected has a radius of 10.351616 ft and covers an area of 336.640351 ft^2.

E5-2

The required spacing for the triangular grid is 20 ft.

Please note that in VSP, the spacing of the triangular grid is not a function of the number of samples that can be obtained using the **Random Start** option.

E5-3

The total cost of hot spot detection for the entire 10 acre site is 10*$ 63,000 = $ 630,000 (for 126 samples).

E5-4

For a fixed criteria of 95% confidence to detect a circular 10-ft radius hot spot:

The only options that can be tried are to use a square or rectangular grid other than the selected triangular grid or to minimize the Measurement Cost per Replicate (i.e. offsite analysis cost).

With the **Random Start** option checked, use of a square grid gives a cost of approximately $54,500 and the use of a rectangular grid gives a cost of $109,000 for one 1-acre site. Compare these costs with $63,000 for one 1-acre site that will be expended using a triangular grid

Minimizing or removing the offsite analysis cost and replacing it with a validated lower field screening cost will provide the best return for the investment made by the contractor.

E5-5

For a fixed criteria of 95% confidence to detect a circular 20-ft radius (at least) hot spot and using a triangular grid (with the length of grid size being 40.0 ft) the total cost of hot spot detection for the entire 10 acre site is 10*$ 16,000 = $160,000. For this example, if the hot spot radius is doubled, the total cost to detect is decreased by a factor of ≈ 4.

E5-6

For a fixed criteria of 95% confidence to detect an ellipse having a semi-major axis of 10-ft., shape of 0.6 and using a triangular grid the total cost of hot spot detection for the entire 10 acre site is 10*$ 128,500 = $1,285,000. For this example, if the hot spot shape is changed to the ellipse, the total cost to detect is increased by a factor of ≈ 2.

Example 6: What if I don't know anything about the site?

E6-1
The number of samples required is 305.

E6-2
The number of samples required is 115.

E6-3
The number of samples required is 73.

E6.4
The number of samples required is 57.

E6-5
The number of samples required is 37.

E6-6
The number of samples required is 27.

E6-7
Yes, if historical data is available, the number of samples required may be reduced because a more realistic (and often lower) estimate of the standard deviation is available. In addition, information regarding the shape of the distribution may also be obtained.