# Probability Distribution

1. In scenario 2, the particle size distribution from the mill is:

| | Counts |
|---|---|
| <10 mm | 50 |
| 11-20 mm | 125 |
| 21-30 mm | 350 |
| 31-40 mm | 275 |
| 41-50 mm | 250 |
| 51-60 mm | 200 |
| 61-70 mm | 40 |
| 71-80 mm | 10 |
| >81 mm | 5 |

Use JMP to perform the following:

(1) Distribution of Counts Vs Size
(2) % Distribution Vs Size
(3) Mean
(4) Variance

2. In scenario 2, the Percent Dissolution of tablets as a function of time is as the following:

| Time | % Dissolution |
|---|---|
| 0 | 0 |
| 15 | 35 |
| 30 | 55 |
| 45 | 70 |
| 60 | 83 |
| 75 | 92 |
| 90 | 97 |
| 105 | 98 |
| 120 | 99 |

Use JMP to plot the Distribution and calculate the time at which 85% of the tablet has been dissolved.

# Comparison Tests

3. Two different catalysts are studied in the batch reactor. (Scenario 1) Differece runs are made with each catalyst and the yield of A measured after 1 hour. (all other factors held constant)

| Catalyst C1 | Catalyst C2 |
|---|---|
| 74 | 71 |
| 70 | 74 |
| 69 | 73 |
| 71 | 75 |
| 72 | 77 |

(1) Determine the mean and variance of each catalyst.
(2) Use the appropriate distribution to decide whether there is a difference at the 95% confidence level.
(3) At what level is there a difference between the two catalyst (p value).
(4) Use an F test to determine the level at which there is a difference between the variance of the yield between the catalysts.

# Regression Analysis

4. Once the API is produced in a reactor described in Scenario 1, crystallization from solution is to separate the desired product $C(t_f)$ from $A(t_f)$ and $B(t_f)$ once the impurity $D(t_f)$ has been removed. In general for a pharmaceutical process crystallization may be used to achieve sufficient product purity, to minimize the filtration time or to achieve tablet stability when mixed with other crystals of other chemical species before forming a tablet. In this example we will dwell only on a single criterion filtration time In this example, based on the work of Togkalidou et al (2001), "Experimental Design and Inferential Modeling in Pharmaceutical Crystallization (AIChe Journal, Vo 27, No1), a pharmaceutical salt was crystallized in a baffled reactor, where the supersaturation was created by adding a less efficient solvent that was miscible in the original solvent. The details are not relevant for the example but the student is referred to the paper if more information about the crystallization process is required.

The following data were collected:

| Experiment Number | Agitation(rpm) | Seed Amount (% of Batch) | Temperature (deg C) | Charge Time h | Filtration Time Min |
|---|---|---|---|---|---|
| 1 | 2200 | 4 | 20 | 6 | 150 |
| 2 | 400 | 5 | 15 | 3 | 105 |
| 3 | 1300 | 3.5 | 15 | 9 | 165 |
| 4 | 2200 | 4 | 17.5 | 7.5 | 170 |
| 5 | 3100 | 3.5 | 17.5 | 7.5 | 90 |
| 6 | 2200 | 4 | 20 | 6 | 155 |
| 7 | 4000 | 5 | 20 | 6 | 50 |
| 8 | 400 | 3 | 20 | 6 | 280 |
| 9 | 1300 | 3.5 | 22.5 | 4.5 | 122 |
| 10 | 2200 | 4 | 22.5 | 4.5 | 100 |
| 11 | 3100 | 4.5 | 25 | 9 | 82 |
| 12 | 2200 | 4 | 20 | 6 | 145 |

Use Regression Analysis from JMP to determine a regression model and the conditions under which the filtration time is minimized.

5. A study was launched to determine the effect of several factors on the %Dissolution after 60 minutes of a new product from the Tabletting machine in Scenario 2. The following data were obtained:

| Expt Number | Speed (Rpm) | Fill Weight (kg) | Pressure (Ton) | Blade Speed (rpm) | Punch Distance (mm) | Powder Flow (kg/hr) | % Diss |
|---|---|---|---|---|---|---|---|
| 1 | 1000 | 100 | 1 | 2000 | 1 | 10 | 50 |
| 2 | 1205 | 110 | .90 | 2010 | .55 | .99 | 77 |
| 3 | 770 | 115 | .91 | 2020 | .48 | .98 | 38 |
| 4 | 750 | 118 | .92 | 2030 | 1.85 | .97 | 83 |
| 5 | 1210 | 120 | .93 | 2040 | 2.05 | .98 | 95 |
| 6 | 820 | 118 | .94 | 2050 | .5 | .99 | 40 |
| 7 | 800 | 115 | .95 | 2060 | 1.9 | .95 | 80 |
| 8 | 1185 | 110 | .96 | 2070 | 2.1 | .98 | 97 |
| 9 | 1200 | 119 | 1.1 | 2080 | .54 | .99 | 75 |
| 10 | 990 | 105 | .97 | 1995 | 1.01 | 10.1 | 55 |
| 11 | 1185 | 95 | 1.4 | 1990 | .52 | 10.2 | 75 |
| 12 | 760 | 85 | 1.5 | 1980 | 2.0 | 10.3 | 69 |
| 13 | 777 | 88 | 1.6 | 1970 | 1.95 | 10.2 | 75 |
| 14 | 1190 | 81 | 1.5 | 1960 | .48 | 10.5 | 80 |
| 15 | 1205 | 105 | 1.3 | 1950 | 2.1 | 10.1 | 98 |
| 16 | 775 | 107 | .95 | 1940 | .52 | 10.6 | 35 |
| 17 | 810 | 75 | 1.2 | 1930 | 2.06 | 10.2 | 60 |
| 18 | 740 | 77 | .97 | 1920 | .47 | 10.1 | 30 |
| 19 | 1010 | 95 | 1.03 | 2010 | .97 | 9.9 | 48 |

a) Determine the extent of correlation between the various factors.
b) Build a regression model relating the %Dissolution to the factors.
    i)Use Standard Regression
    ii)Use Stepwise Regression
    iii) Why are results in ii) different than in i)

# Single Factor Experiments

## 6. Completely Randomized Design

In a study to determine the effect of roller speed on roller gap in a roller compactor (Scenario 2), five replicates of the Roller Gap in mm were measured at five different values of roll speed (rpm) where the experiments were run in random order. The following data were obtained:

| Roll Speed (rpm) | Roller gap (mm) | | | | |
|---|---|---|---|---|---|
| 15 | 7 | 7 | 15 | 11 | 9 |
| 20 | 12 | 17 | 12 | 18 | 18 |
| 25 | 14 | 18 | 18 | 19 | 19 |
| 30 | 19 | 25 | 22 | 19 | 23 |
| 35 | 7 | 10 | 11 | 15 | 11 |

(1) Does roller speed affect roller gap at the 95% confidence level? Perform an ANOVA.

(2) Using a multiple range test at 95% confidence which levels are different from one another?

(3) Find a suitable regression model between roller gap and roll speed if one exists.

(4) Compare the results of (2) and (3).

# Randomized Block Design

7. A study was conducted to determine effect of Roll Speed (rpm) on ribbon uniformity (dimensionless) in a roller compactor (Scenario 2).. Six different replicates were conducted on six batches of material from a blending operation. The order of selecting the samples was from the blenders were randomized as was the order of running the experiments. The data from this completely randomized block design is shown below:

| | Batch Number | | | | | |
|---|---|---|---|---|---|---|
| Roll Speed (rpm) | 1 | 2 | 3 | 4 | 5 | 6 |
| 10 | .78 | .80 | .81 | .75 | .77 | .78 |
| 16 | .85 | .85 | .92 | .86 | .81 | .83 |
| 23 | .93 | .92 | .95 | .89 | .89 | .83 |
| 31 | 1.14 | .97 | .98 | .88 | .86 | .83 |
| 40 | .97 | .86 | .78 | .76 | .76 | .75 |

a.) Does Roll Speed affect the ribbon uniformity? Is the between batch variation significant?
b.) Determine the regression equation between roller uniformity and roll speed. Compare the results with a)
c.) Are the residuals from this experiment normally distributed?

# Optimization Problem

8. The product uniformity y from a continuous blender in scenario 2 is related to the tilt(deg) T by the relationship:

$$Y = \begin{cases} 100-(20.5-T)^2+ \varepsilon, & \text{if } Y>0 \\ 0, & \text{if } Y\leq 0 \end{cases}$$

It is clear from the above relationship that the maximum uniformity is obtained at T=20.5

Show how (1) dichotomous search and (2) golden section search can be used to search out this optimum over the region $0\leq T\leq 50$ where the measurement error at any point is

$\varepsilon \sim N(0,.25)$

The smallest difference in T which can be detected is 2 degree.

(Hint: Program the relationship in Excel using the available random number generator)

# Factorial Experimentation

9. A study is conducted to assess the effect of Pressure (Ton) and Punch Distance (mm) on percent dissolution of a new API after 80 minutes in a Tablet Press in Scenario 2. Three different replicates were taken at random at three pressures and two Punch Distances The data are as follows:

Pressure (Ton)

| Punch Distance (mm) | .75 | 1 | 1.5 |
|---|---|---|---|
| 1 | 74,64,50 | 73,61,44 | 78,85,92 |
| 2 | 92,86,68 | 98,73,88 | 66,45,85 |

(1) Build a mathematical model to describe the mathematical relationship between %Dissolution and (Pressure, Punch Distance).

(2) Analyze the residuals from this experiment.

10. Design a full factorial experiment to determine the effect of Tilt, Speed, Load and Inlet powder flow on the uniformity and density in a series of batch runs in a continuous blender in scenario 2. Consider the following cases:

(a) All factors at two levels.

(b) All factors at three levels.

(c) Tilt at 2 levels, Speed at three levels, load at four levels and inlet powder flow at 2 levels.

(1) For each of these cases give the following:

i) the actual experiments that must be run.
ii) the mathematical model

(2) Describe the role of replication, randomization and blocking

# Fractional Factorial Experiments with two levels

11. In the investigation of the conditions of filtration during the preparation of an API, the objective was to improve the quality of the product. Four factors were examined:

A. Concentration of liquor when filtered (concentrated v. dilute)

B. Effect of Liquor Storage (fresh vs old). The liquor was either filtered immediately or kept a week before filtration.

C. Presence or absence of an anti-frothing agent.

D. Temperature of Filtration (high vs low)

It was considered unlikely that large interactions would exist between these factors so that a ½ replicate of a $2^4$ factorial was selected with defining contrast D=ABC. The purity of the product was recorded in the table below:

| Run No.. | A | B | C | D | Purity |
|---|---|---|---|---|---|
| 1 | -1 | -1 | -1 | -1 | 107 |
| 2 | 1 | -1 | -1 | +1 | 114 |
| 3 | -1 | 1 | -1 | 1 | 122 |
| 4 | +1 | +1 | -1 | -1 | 130 |
| 5 | -1 | -1 | 1 | 1 | 106 |
| 6 | 1 | -1 | +1 | -1 | 121 |
| 7 | -1 | +1 | +1 | -1 | 120 |
| 8 | 1 | 1 | 1 | 1 | 132 |

Determine:

(1) The pattern of aliases for the experiment.

(2) The main effects and interactions

(3) If the error in the measurements is 2 units, which factors are significant?

12. O.L. Davies. The following experiment was conducted in a batch reactor (Scenario 1) to investigate conditions affecting the yield of an API. Five factors were investigated with the following levels:

| Factors | Level | |
|---|---|---|
| A A/B Feed ratio | Low | High |
| B Amount of Liquid Catalyst | Concentrated | Dilute |
| C Amount of Anti-foaming agent | None | Some |
| D Time of Reaction | Short | Fast |
| E Agitation | Slow | Fast |

Setting the signs of D=-AE and C=+AB, the following Percent Yield data were obtained (the analysis for each run was repeated)

Design of Experiment and Product Yield

| Run No | A | B | C | D | E | Yield |
|---|---|---|---|---|---|---|
| 1 | -1 | -1 | +1 | -1 | -1 | 53.1,54.6 |
| 2 | +1 | -1 | -1 | +1 | -1 | 49.3,48.4 |
| 3 | -1 | +1 | -1 | -1 | -1 | 50.1,51.4 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4 | +1 | +1 | +1 | +1 | -1 | 68.3,67.4 |
| 5 | -1 | -1 | +1 | +1 | +1 | 73.4,75.3 |
| 6 | +1 | -1 | -1 | -1 | +1 | 79.7,78.0 |
| 7 | -1 | +1 | -1 | +1 | +1 | 84.5,86.4 |
| 8 | +1 | +1 | +1 | -1 | +1 | 81.3,80.4 |

(1) What are the defining contrasts?

(2) Determine the pattern of aliases.

(3) What are the significant main effects and interactions?

(4) Is there a significant lack of fit?

(5) Based on this data what is the optimal way to run the reaction?

13. In the batch reaction API yield study described in scenario 1, it was decided to make a series of runs including temperate as well as the other five factors. Based on their previous success they were allowed to conduct 16 runs.

(1) Design a fractional factorial experiment which is a ¼ fraction of a $2^6$ full factorial experiment which maximizes the probability of testing for the significant of main effect and two factor interactions.

(2) What are the defining contrasts and pattern of aliases for this design.

(3) List the considerations in deciding which fraction to run.

# Response Surface Modeling and Optimization

14. An experiment was run in a batch reactor to determine the effect of temperature and reaction time on the yield of the API. These factors are coded as x1= (temperature -300deg)/50deg and x2=(time-10hrs)/5 hours. The following coded data was obtained where the yield is in percent

| Run No | X1 | X2 | Yield (%) |
|---|---|---|---|
| 1 | -1 | 0 | 78.03 |
| 2 | 1 | 0 | 80.4 |
| 3 | 0 | 0 | 80.1 |
| 4 | 0 | 0 | 80.95 |
| 5 | 0 | -1 | 80.3 |
| 6 | 0 | 1 | 80.08 |
| 7 | 0 | 0 | 80.97 |
| 8 | -1.4142 | -1.4142 | 74.38 |
| 9 | -1.4142 | 1.4142 | 74.87 |
| 10 | 1.4142 | -1.4142 | 75.68 |
| 11 | 1.4142 | 1.4142 | 78.13 |
| 12 | 0 | 0 | 80.44 |

(1) Fit a response surface model to the data. Is the model adequate to describe the data?

(2) Plot the yield response curve. What recommendations would you make about the operating conditions for the reactor?

15. Design a Central Composite Design, a Three Level Factorial Design and a Box Behnken design for generating a response surface for yield in a batch reactor system(Scenario 1) where the effect of temperature, termination time and agitation rate are to be investigated. Compare the features of the three designs in terms of the number of runs required.