Categoric Mixture Components Proportion Going to Zero

Pat Whitcomb
Stat-Ease, Inc.
2021 E. Hennepin Ave
Suite 480
Minneapolis, MN 55413
Ph 612.746.2036
fax 612.746.2056
pat@statease.com

Stat-Ease, Inc.
2021 E. Hennepin Ave
Suite 480
Minneapolis, MN 55413

2012 Fall Technical Conference

Agenda

- **Categoric Mixture Components Proportion Going to Zero**
  - **What's the problem?**
  - Shelf life example
  - Conclusion

Categoric Factor with Proportion Zero
Four Component and One Categoric

Let's start with a four component example where (D) is proportion of one, or the other, of two categoric components:

- A, B, C and D vary from 0 to 1
- D is the proportion of categoric factor e, which has two levels (e_1 and e_2)

Build a combined design for a quadratic mixture crossed with a main effect model for the categoric factor (e):

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD$$
$$+ \beta_{15} Ae + \beta_{25} Be + \beta_{35} Ce + \beta_{45} De + \beta_{125} ABe + \beta_{135} ACe + \beta_{145} ADe$$
$$+ \beta_{235} BCe + \beta_{245} BDe + \beta_{345} CDe$$

Categoric Factor with Proportion Zero

Four Component and One Categoric

Component 1 A:A mg	Component 2 B:B mg	Component 3 C:C mg	Component 4 D:D mg	Factor 5 E:E
100.000	1000.000	0.000	0.000	e2
320.392	728.066	51.542	0.000	e2
[illegible]	066	51.542	0.000	e2
[illegible]	456	52.695	0.000	e1
[illegible]	223	120.000	0.000	e2
[illegible]	456	52.695	0.000	e1
[illegible]	000	112.000	0.000	e2
485.374	500.000	114.626	0.000	e2
292.094	687.906	120.000	0.000	e1
493.681	606.319	0.000	0.000	e1
600.000	500.000	0.000	0.000	e2

Problem: When the proportion of a categoric factor goes to zero, then the different levels are all the same – **absent!**

Categoric Factor with Proportion Zero

Four Component and One Categoric

Problem: When the proportion of a categoric factor goes to zero, then the different levels are all the same – absent!

Four component example:

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD$$

$$+ \beta_{15} Ae + \beta_{25} Be + \beta_{35} Ce + \beta_{45} De + \beta_{125} ABe + \beta_{135} ACe + \beta_{145} ADe$$

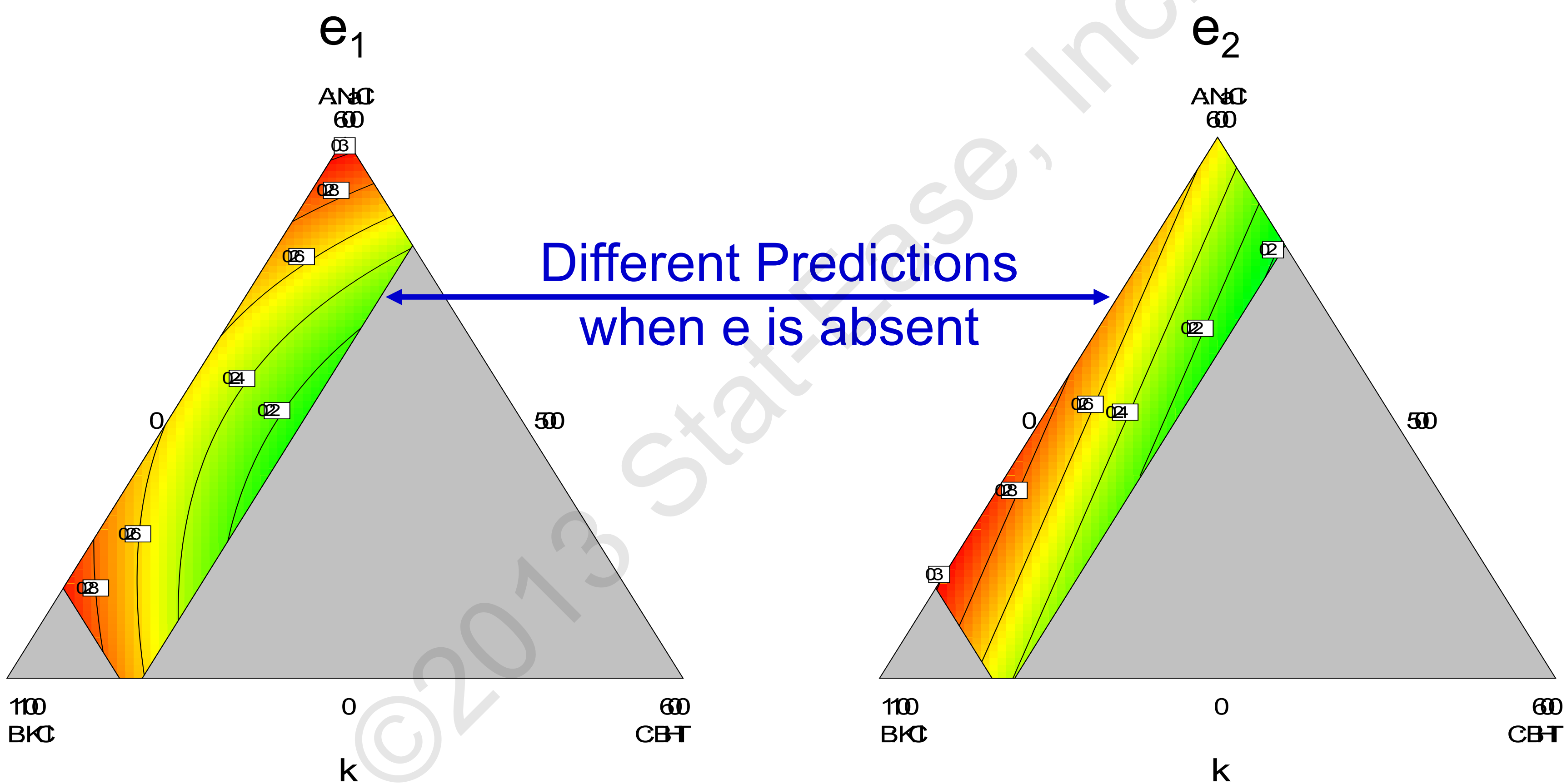
$$+ \beta_{235} BCe + \beta_{245} BDe + \beta_{345} CDe$$

Predictions can vary for each level of the categoric component even when the categoric component is absent (D=0)!

(see next slide)

Categoric Factor Absent, i.e. D = 0

Different Predictions for each Levels



Categoric Factor with Proportion Zero
Four Component and One Categoric

Solution: Start with the mixture model and add:

- The mixture terms with "D" crossed with "e".
- The mixture terms without "D" crossed with "De".

Never let "e" appear without "D".

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD$$
$$+ \beta_{45} De + \beta_{145} ADe + \beta_{245} BDe + \beta_{345} CDe + \beta_{1245} ABDe + \beta_{1345} ACDe$$
$$+ \beta_{2345} BCDe$$

Now the predictions are the same for each level of the categoric component when the categoric component is absent (D=0)!

(see next slide)

Mixture model
+ the mixture terms with "D" crossed with "e"
+ the mixture terms without "D" crossed with "De":

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD$$
$$\left(+\beta_4 D + \beta_{14} AD + \beta_{24} BD + \beta_{34} CD\right)(e)$$
$$\left(+\beta_1 A + \beta_2 B + \beta_3 C + \beta_{12} AB + \beta_{13} AC + \beta_{23} BC\right)(De)$$

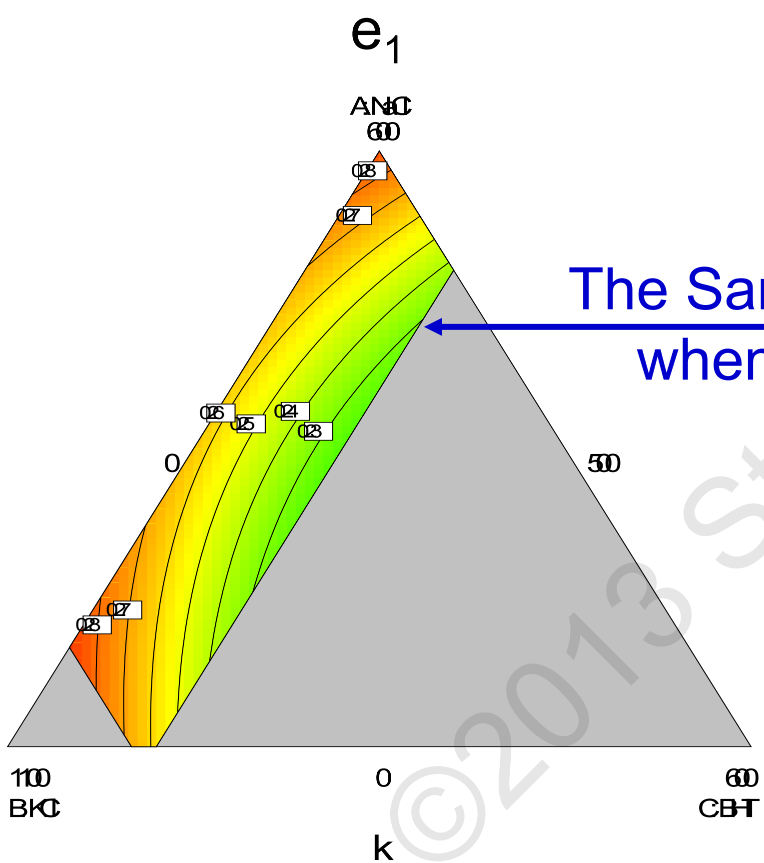
$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD$$
$$+ \beta_{45} De + \beta_{145} ADe + \beta_{245} BDe + \beta_{345} CDe$$
$$+ \beta_{145} ADe + \beta_{245} BDe + \beta_{345} CDe + \beta_{1245} ABDe + \beta_{1345} ACDe + \beta_{2345} BCDe$$

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD$$
$$+ \beta_{45} De + \beta_{145} ADe + \beta_{245} BDe + \beta_{345} CDe + \beta_{1245} ABDe + \beta_{1345} ACDe + \beta_{2345} BCDe$$

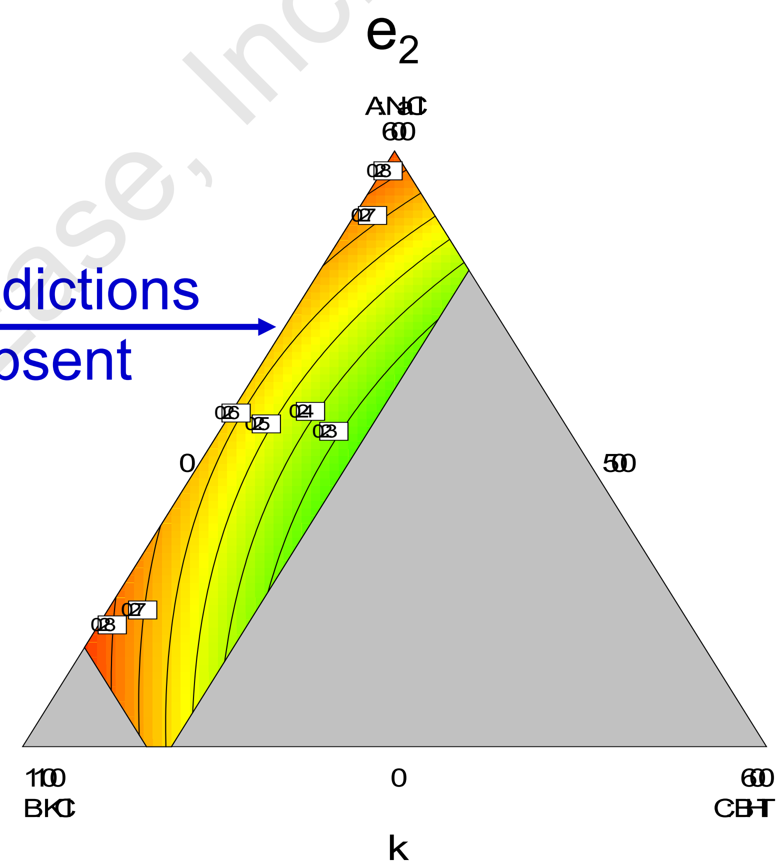
"e" never appears without "D"

Categoric Factor Absent, i.e. D = 0

Same Predictions for each Levels



The Same Predictions
when e is absent



Categoric Factor with Proportion Zero

Mixture Quadratic with One Categoric

x_q is the proportion of a component and z_1 is the levels (categoric level) of that component :

average blending properties | correction to average blending for categoric amount and levels

$$\eta(x,z) = \sum_{i=1}^{q} \gamma_i^0 x_i + \sum_{i<} \sum_{j}^{q} \gamma_{ij}^0 x_i x_j + \gamma_1^0 x_q z_1 + \left[\sum_{i=1}^{q-1} \gamma_i^1 x_i + \sum_{i=1<} \sum_{j}^{q-1} \gamma_{ij}^1 x_i x_j \right] (x_q z_1)$$

treat $x_q z_1$ as one term (amount)(levels) of a component

q = number of mixture components
m = number of mixture components with categoric levels
(in this case m is fixed at one: m =1)

Categoric Factor with Proportion Zero

Mixture Quadratic with One Categoric

average blending properties

correction to average blending for categoric amount and levels

$$\eta(x,z) = \sum_{i=1}^{q}\gamma_i^0 x_i + \sum_{i<}\sum_{j}^{q}\gamma_{ij}^0 x_i x_j + \gamma_1^0 x_q z_1 + \left[\sum_{i=1}^{q-1}\gamma_i^1 x_i + \sum_{i=1<}\sum_{j}^{q-1}\gamma_{ij}^1 x_i x_j\right](x_q z_1)$$

for q=4, m=1

$$\hat{y} = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

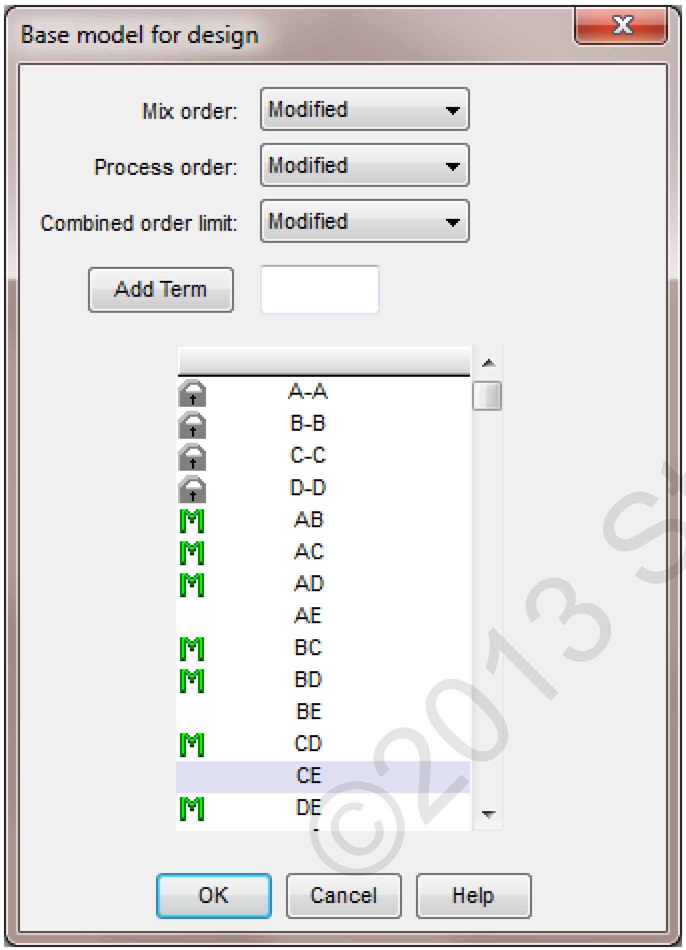
$$+\beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{14}x_1x_4 + \beta_{23}x_2x_3 + \beta_{24}x_2x_4 + \beta_{34}x_3x_4$$

$$+\beta_{41}x_4z_1 + \beta_{141}x_1x_4z_1 + \beta_{241}x_2x_4z_1 + \beta_{341}x_3x_4z_1$$

$$+\beta_{1241}x_1x_2x_4z_1 + \beta_{1341}x_1x_3x_4z_1 + \beta_{2341}x_2x_3x_4z_1$$

Categoric Factor with Proportion Zero

Four Component and One Categoric



model terms (q=4, m=1):

A, B, C, D,

AB, AC, AD, BC, BD, CD,

DE,

ADE, BDE, CDE,

ABDE, ACDE, BCDE

Categoric Factor with Proportion Zero
Four Component and One Categoric

The coupled model we specify while building the design is called the "Design Model":

A, B, C, D, AB, AC, AD, BC, BD, CD,
DE, ADE, BDE, CDE, ABDE, ACDE, BCDE.

Analyzing a response, start with the "Design Model" and reduce it.

It is important to note the design model is not hierarchical, e.g. for ADE there is no AE. Therefore we say "NO" to Design-Expert's query "Would you like the hierarchy corrected automatically?"

We must check hierarchy ourselves. E.g. if ABDE is selected be sure to include AB, AD, BD, DE, ADE and BDE, but not AE or BE.

When D = 0
then the equations are different

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD$$
$$+ \beta_{15} Ae + \beta_{25} Be + \beta_{35} Ce + \beta_{45} De + \beta_{125} ABe + \beta_{135} ACe + \beta_{145} ADe$$
$$+ \beta_{235} BCe + \beta_{245} BDe + \beta_{345} CDe$$

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_{12} AB + \beta_{13} AC + \beta_{23} BC + \beta_{15} Ae + \beta_{25} Be + \beta_{35} Ce$$
$$+ \beta_{125} ABe + \beta_{135} ACe + \beta_{235} BCe$$

D is zero

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD + \beta_{45} De$$
$$+ \beta_{145} ADe + \beta_{245} BDe + \beta_{345} CDe + \beta_{1245} ABDe + \beta_{1345} ACDe + \beta_{2345} BCDe$$

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_{12} AB + \beta_{13} AC + \beta_{23} BC$$

D is zero

When D > 0 then the equations are equivalent

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD$$
$$+ \beta_{15} Ae + \beta_{25} Be + \beta_{35} Ce + \beta_{45} De + \beta_{125} ABe + \beta_{135} ACe + \beta_{145} ADe$$
$$+ \beta_{235} BCe + \beta_{245} BDe + \beta_{345} CDe$$

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_{12} AB + \beta_{13} AC + \beta_{23} BC + \beta_{45} e$$
$$+ \beta_{15} Ae + \beta_{25} Be + \beta_{35} Ce + \beta_{125} ABe + \beta_{135} ACe + \beta_{235} BCe$$

D is a fixed value > 0

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4 D + \beta_{12} AB + \beta_{13} AC + \beta_{14} AD + \beta_{23} BC + \beta_{24} BD + \beta_{34} CD + \beta_{45} De$$
$$+ \beta_{145} ADe + \beta_{245} BDe + \beta_{345} CDe + \beta_{1245} ABDe + \beta_{1345} ACDe + \beta_{2345} BCDe$$

$$\hat{y} = \beta_1 A + \beta_2 B + \beta_3 C + \beta_{12} AB + \beta_{13} AC + \beta_{23} BC + \beta_{45} e$$
$$+ \beta_{145} Ae + \beta_{245} Be + \beta_{345} Ce + \beta_{1245} ABe + \beta_{1345} ACe + \beta_{2345} BCe$$

D is a fixed value > 0

Categoric Factor with Proportion Zero
Four Component and One Categoric

Since:

- The different levels of a component are represented as a categoric factor.
- Actual proportions of the categoric components are key.
- Strict hierarchy is not maintained.

The analysis should be done using "**Real**" coding.

Agenda Transition

- **Categoric Mixture Components Proportion Going to Zero**
  - What's the problem?
  - **Shelf life example**
  - Conclusion

Shelf Life Characterization
Categoric Factor Going to Zero

Optimize a preservative blend to maximize shelf life of a new food:

- One of two additional preservatives (either sulfur dioxide or calcium propionate) will be used.

Component	Range
A: NaCl	0 – 600 mg
B: KCl	0 – 500 mg
C: BHT	0 – 120 mg
D: Preservative	0 – 120 mg
Total =	1100 mg

- MLC: the sum of the antioxidant BHT (butylated hydroxytoluene) and Preservative must be ≤ 130 mg.

Shelf Life Characterization
Degradation Rate

Degradation is a function of time:

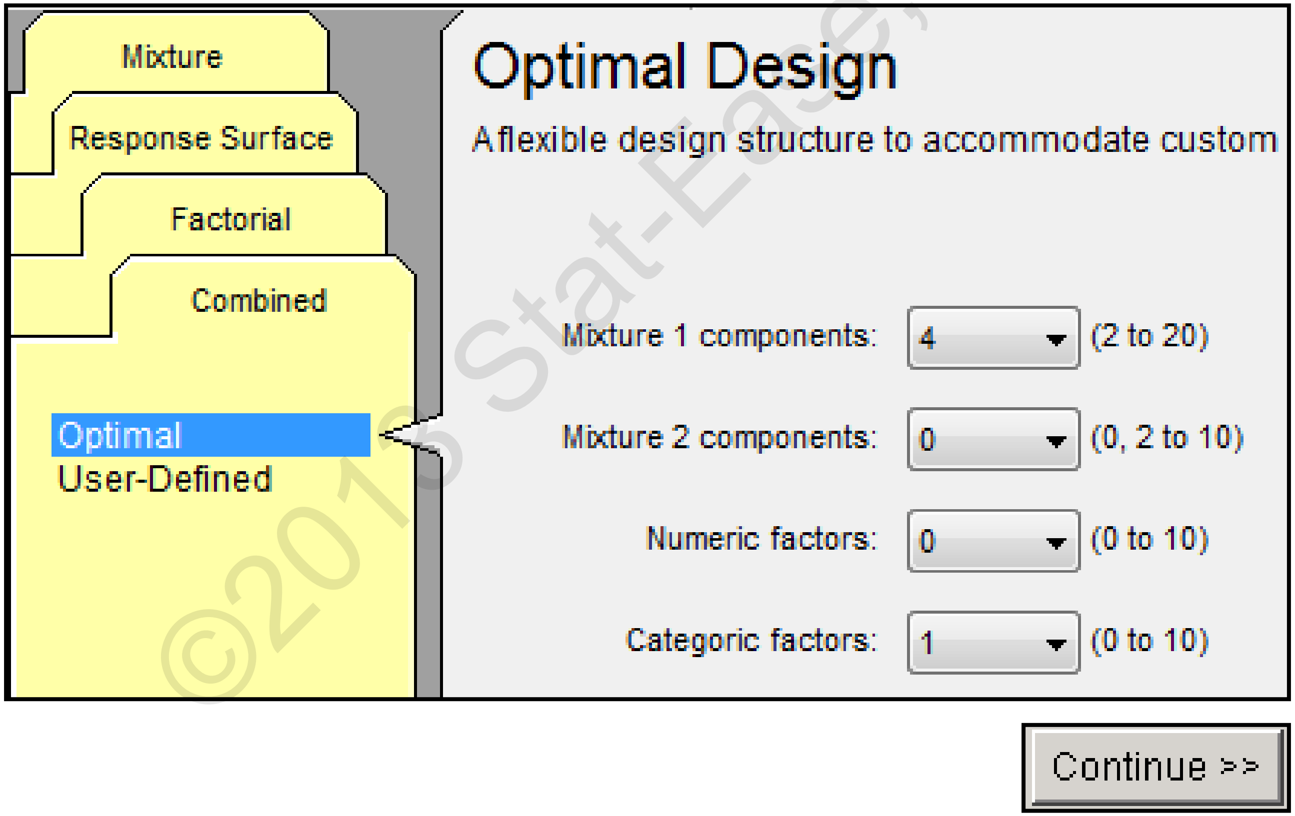
- It is assumed that degradation rate is zero order because of a low concentration of degradation products.
- If degradation is zero order, then a plot of the amount of degradation product versus time is linear with a slope of k.
- The major degradation was measured over time using accelerated life testing for each of the 27 runs and the rate (k) determined for each mixture combination.

Lower rates (k values) are better!

Categoric Factor Going to Zero

Shelf Life Characterization: Build *(page 1 of 6)*

1. Build an "**Optimal**" **Combined** design with four components and one categorical factor:



Categoric Factor Going to Zero

Shelf Life Characterization: Build *(page 2 of 6)*

Stat-Ease statistics made easy®

2. Enter mixture components:

Mix Components: 4

Total: 1100

Units: mg

(•) Horizontal ( ) Vertical

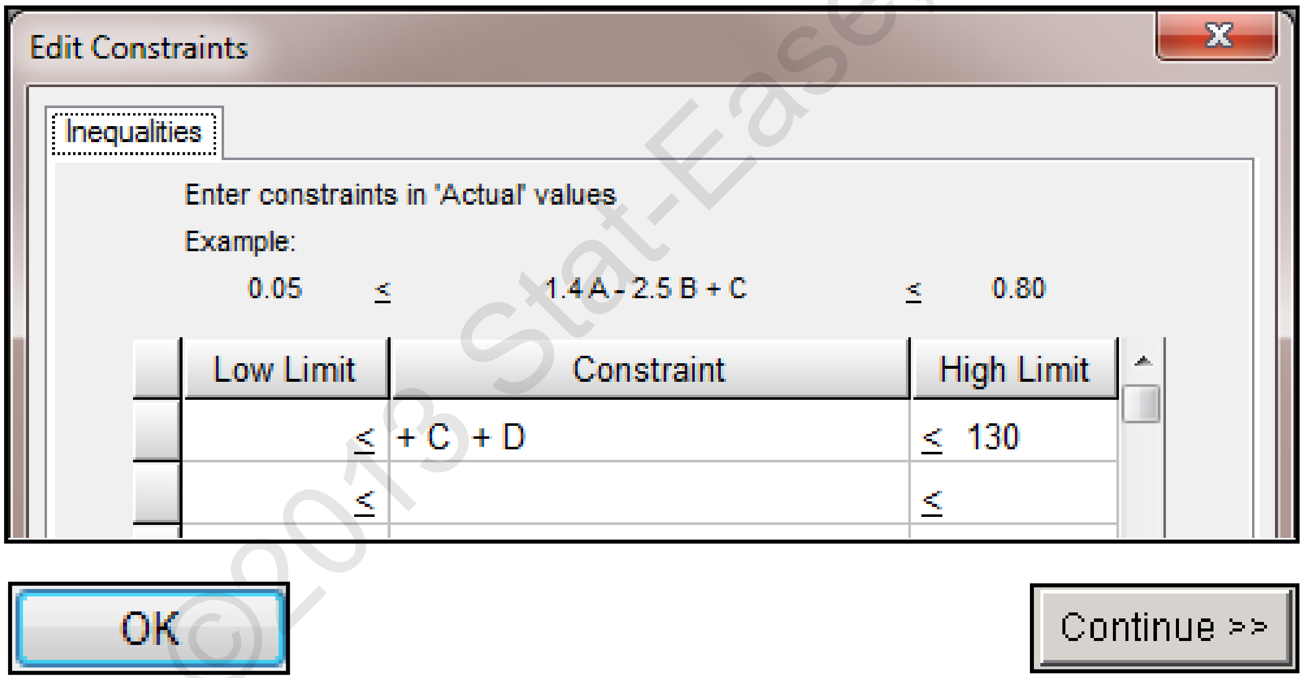
	Name	Low	High
A [Mixture]	NaCl	0	600
B [Mixture]	KCl	500	1000
C [Mixture]	BHT	0	120
D [Mixture]	Preservative	0	120

enter the MLC constraint on the next slide.

Categoric Factor Going to Zero

Shelf Life Characterization: Build *(page 3 of 6)*

3. Click on the "Edit constraints…" button and enter the constraint for the maximum of C (BHT) + D(preservative) ≤ 130 mg :



Categoric Factor Going to Zero
Shelf Life Characterization: Build *(page 4 of 6)*

4. Enter categoric factors:

Numeric Factors: 0 — Horizontal

Categoric factors: 1 — Vertical (selected)

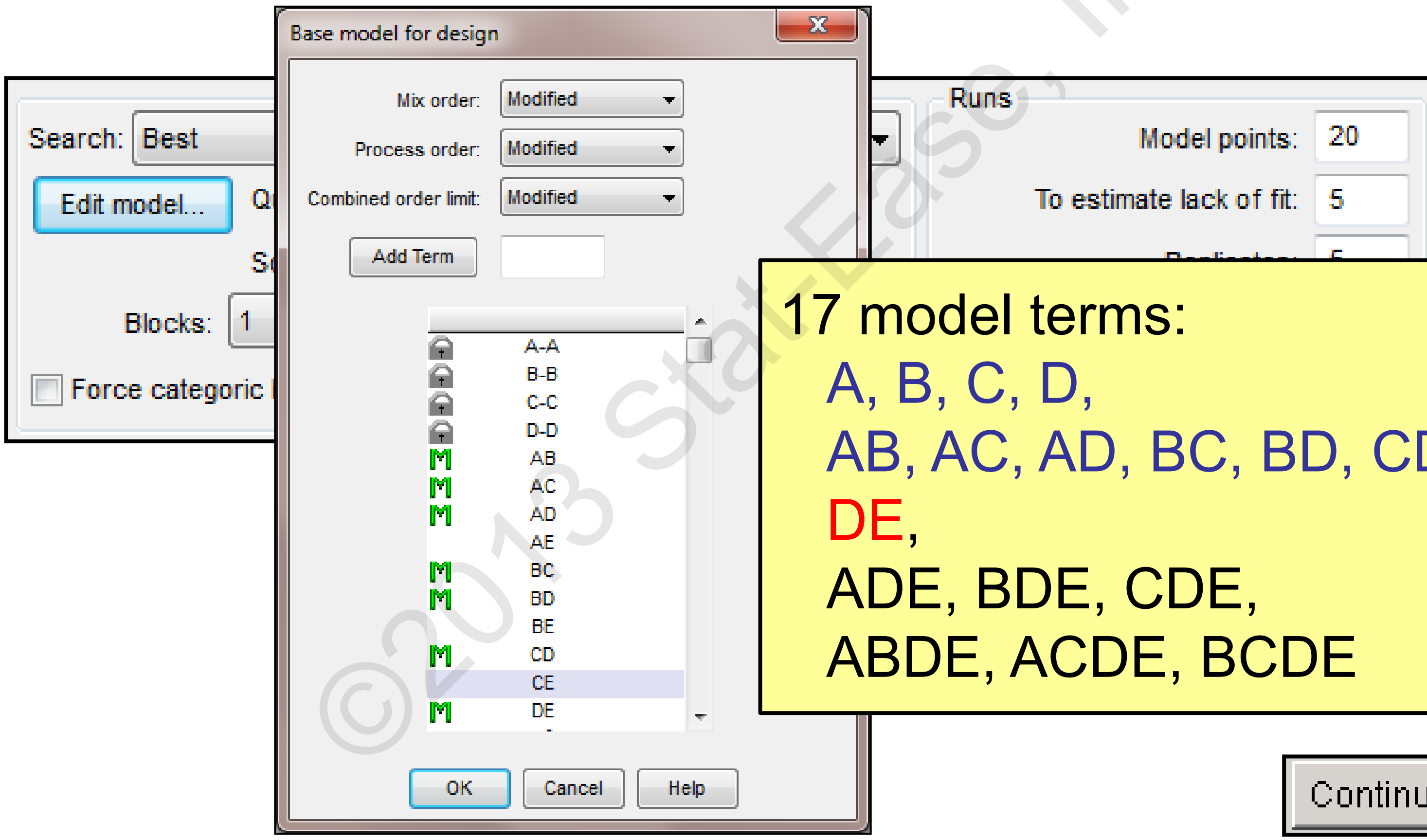
	E [Categoric]
Name	Species
Units	preservative
Type	Nominal
Levels	2
L[1]	S dioxide
L[2]	Ca propionate

Continue >>

Categoric Factor Going to Zero

Shelf Life Characterization: Build *(page 5 of 6)*

5. Click "**Edit model…**" and design for a "going to zero" model:



17 model terms:

A, B, C, D,

AB, AC, AD, BC, BD, CD,

DE,

ADE, BDE, CDE,

ABDE, ACDE, BCDE

Continue >>

Categoric Factor Going to Zero
Shelf Life Characterization: Build *(page 6 of 6)*

6. There is one response:

Name	Units
k	rate constant

Continue >>

File Edit View Display Options Design Tools Help Tips

- Notes for Categoric mix co
  - Design (Actual x Actu
    - Summary
    - Graph Columns
    - Evaluation
    - Constraints
  - Analysis
    - R1:k
  - Optimization
    - Numerical
    - Graphical
    - Point Prediction
    - Confirmation

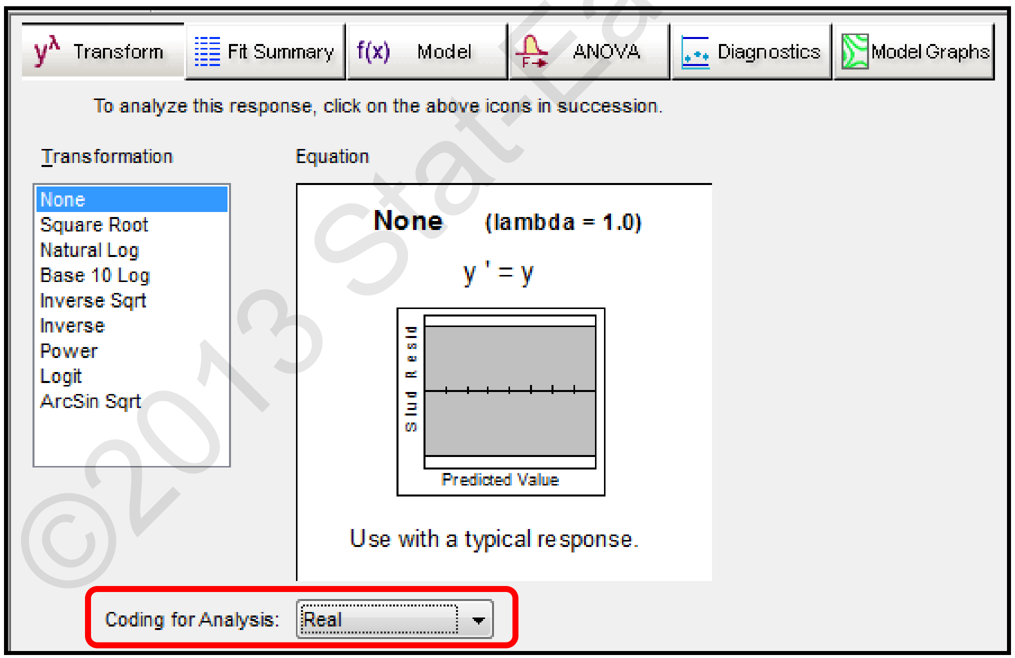
Std	Run	Component 1 A:NaCl mg	Component 2 B:KCl mg	Component 3 C:BHT mg	Component 4 D:Preservative mg	Factor 5 E:Species preservative	Response 1 k rate constant
4	1	480.000	500.000	120.000	0.000	S dioxide	
14	2	0.000	1000.000	0.000	100.000	S dioxide	
20	3	478.188	524.867	45.822	51.123	Ca propionate	
11	4	348.727	681.427	10.960	58.886	S dioxide	
25	5	199.273	770.727	63.668	66.332	Ca propionate	
19	6	317.774	717.474	64.752	0.000	Ca propionate	
17	7	561.344	538.656	0.000	0.000	Ca propionate	
2	8	268.502	831.498	0.000	0.000	S dioxide	
26	9	454.621	525.379	0.000	120.000	Ca propionate	

Categoric Factor Going to Zero

Shelf Life Characterization: Analysis *(page 1 of 2)*

1. Right click on the response column header and "Simulate Response" using "**degradation rate.sim**".

2. Remember: the analysis should be done use "**Real**" coding:



Categoric Factor Going to Zero

Shelf Life Characterization: Analysis *(page 2 of 2)*

3. Fit an appropriate model to the response.
   This requires skipping the fit summary and using the "**Design Model**".
   *(Say "**No**" to hierarchy – see notes below.)*

f(x) Model

4. Recommend a formula to minimize degradation rate (k):

NaCl ______ mg

KCl ______ mg

BHT ______ mg

Preservative ______ mg of "Sulfur dioxide" or "Calcium propionate"

(circle one)

Agenda Transition

- **Categoric Mixture Components Proportion Going to Zero**
  - What's the problem?
  - Shelf life example
  - **Conclusion**

Multiple Categoric Factors Going to Zero (without categoric interaction)

$x_{q\text{-}m+k}$ is the proportion of a component and z_k is the levels (categoric level) of that component:

$$\eta(x,z)=\sum_{i=1}^{q}\gamma_i^0 x_i+\sum_{i<}\sum_{j}^{q}\gamma_{ij}^0 x_i x_j+\sum_{k=1}^{m}\left[\gamma_k^0 x_{q-m+k} z_k+\left[\sum_{i=1}^{q-m}\gamma_i^1 x_i+\sum_{i=1<}\sum_{j}^{q-m}\gamma_{ij}^1 x_i x_j\right]\left(x_{q-m+k} z_k\right)\right]$$

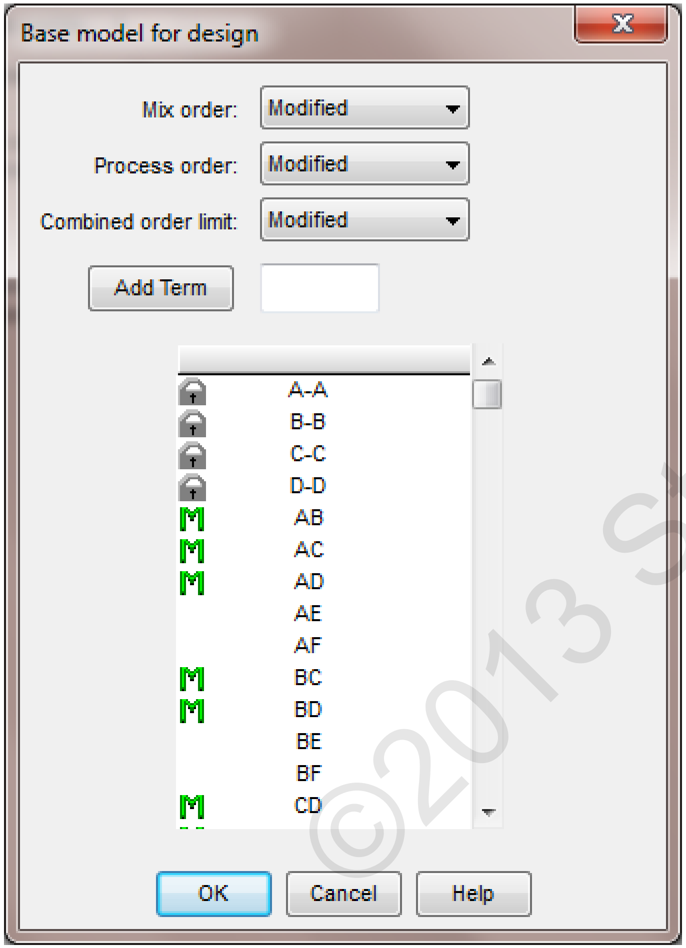
q – number of components

m – number of components with different levels

Model terms for q=4 and m=2:

A, B, C, D, AB, AC, AD, BC, BD, CD,
CE, ACE, BCE, ABCE, DF, ADF, BDF, ABDF

Two Components with Categoric Factors Going to Zero



model terms (q=4, m=2):

A, B, C, D,
AB, AC, AD, BC, BD, CD,
CE, DF,
ACE, BCE, ABCE,
ADF, BDF, ABDF

Multiple Categoric Factors Going to Zero (with categoric interaction)

x_{q-m+k} is the proportion of a component and z_k is the levels (categoric level) of that component:

$$\eta(x,z) = \sum_{i=1}^{q}\gamma_i^0 x_i + \sum_{i<}\sum_{j}^{q}\gamma_{ij}^0 x_i x_j + \sum_{k=1}^{m}\left[\gamma_k^0 x_{q-m+k} z_k + \left[\sum_{i\neq k}^{q}\gamma_i^1 x_i + \sum_{i\neq k<}\sum_{j\neq k}^{q}\gamma_{ij}^1 x_i x_j\right]\left(x_{q-m+k} z_k\right)\right]$$

$$+\sum_{k}^{m-1}\left(x_{q-m+k} z_k\right)\left(x_{q-m+k+1} z_{k+1}\right)$$

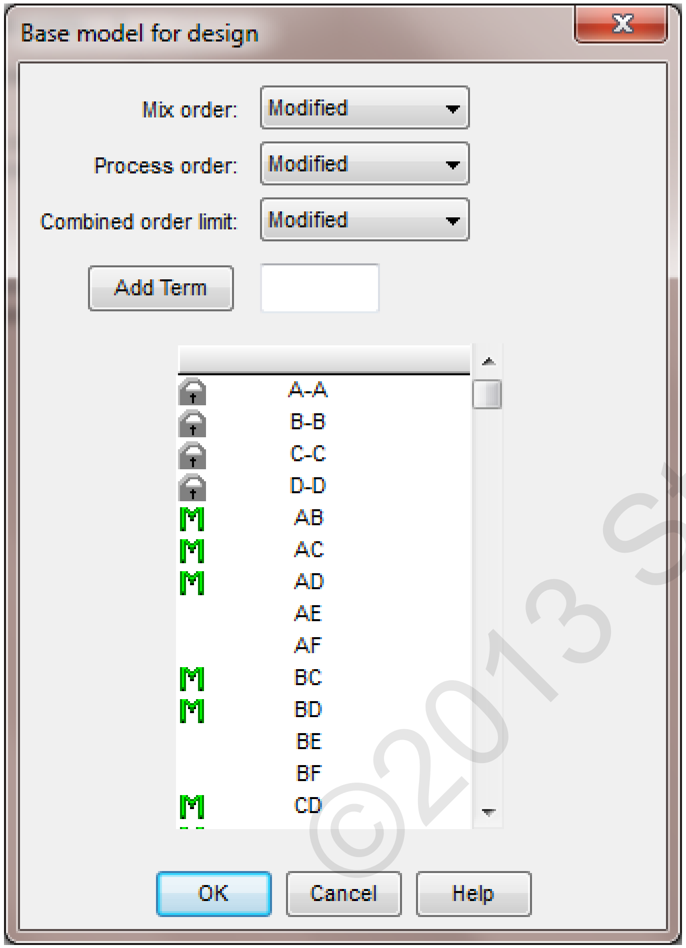
q – number of components

m – number of components with different levels

Model terms for q=4 and m=2:

A, B, C, D, AB, AC, AD, BC, BD, CD,
CE, ACE, BCE, DCE, ABCE, ADCE, BDCE,
DF, ADF, BDF, CDF, ABDF, ACDF, BCDF, CEDF

Two Components with Categoric Factors Going to Zero



Quadratic model terms (q=4, m=2):

A, B, C, D,
AB, AC, AD, BC, BD, CD,
CE, ACE, BCE, CDE,
ABCE, ACDE, BCDE,
DF, ADF, BDF, CDF,
ABDF, ACDF, BCDF,
CDEF