USING INNOVATIVE STATISTICS FOR INCREASED QUALITY, FASTER DEVELOPMENT

FALL TECHNICAL CONFERENCE 2017

KATHERINE E.D. GIACOLETTI, PARTNER, SYNOLOSTATS

SynoloStats

INTEGRATING TOTAL SOLUTIONS

OUTLINE

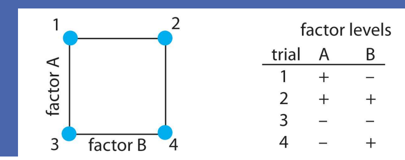
- Why use statistics?
- Why use new statistical methods?
- Motivating examples
  - Innovations in DOE
  - Innovations in prediction

WHY USE STATISTICS?

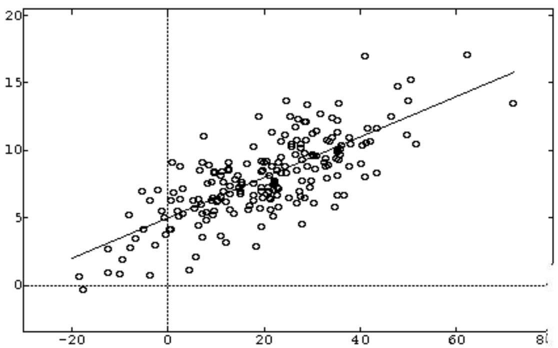
- Understand variability
- Understand processes (output $\propto$ inputs)
- Characterize performance
- Detect unusual performance

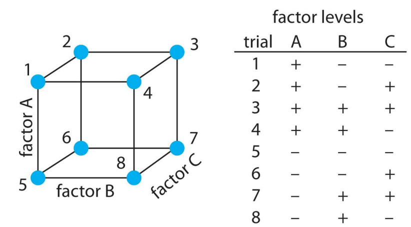
quantify & reduce risk

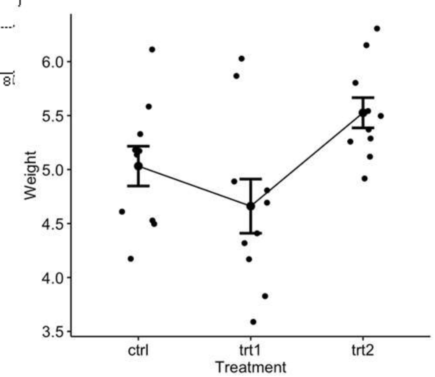
STATISTICS FOR RISK MANAGEMENT



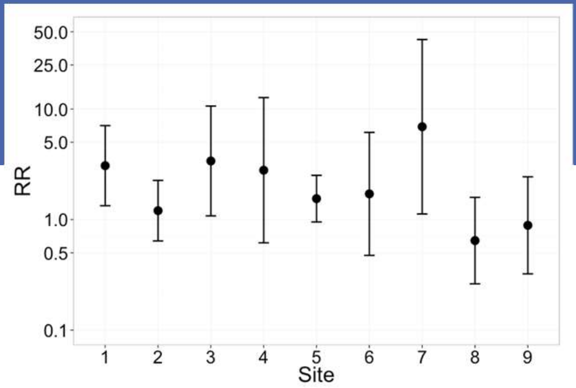
- Understanding variability & processes
  - Design of Experiments (factorials, fractional factorials, response surface, etc.)
  - Regression, ANOVA, variance components



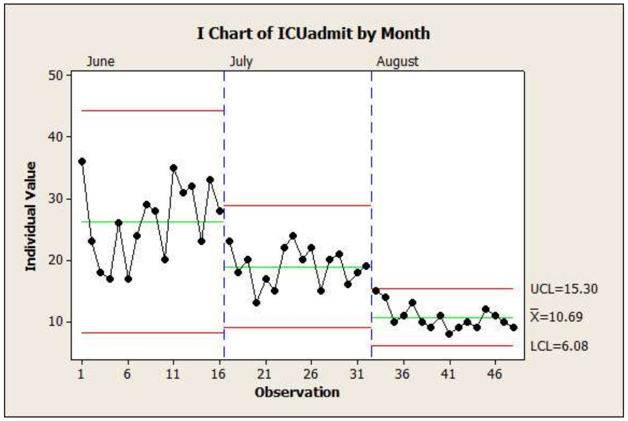




STATISTICS FOR RISK MANAGEMENT



- Characterize performance
  - Confidence, prediction, tolerance intervals
  - Capability indices, control limits
- Detect unusual performance
  - Control charts



WHY USE NEW STATISTICAL METHODS?

- All the same reasons, but why not do it *better?*
  - More efficiently
    - Modern DOE methods, incorporate prior knowledge in analysis
  - More clearly
    - Probability rather than confidence
  - More precisely
    - Predictions incorporating all sources of variability

ADVANCES IN EXPERIMENTAL DESIGN

CLASSICAL DESIGN OF EXPERIMENTS

- 1940s Fractional Factorial, Placket Burman
- 1970s Optimal Design (theory much older)
- 2011 Definitive Screening Designs

CLASSICAL – SCREENING DESIGNS: ALIASING OF EFFECTS

Effects that cannot be estimated separately from each other are said to be confounded

Resolution III

No main effects are aliased with any other main effect, but main effects are aliased with 2-factor interactions.

Resolution IV

No main effects are aliased with any other main effect or 2-factor interactions, but some 2-factor interactions are aliased with other 2-factor interactions and main effects are aliased with 3-factor interactions.

Resolution V

No main effects or 2-factor interactions are aliased with any other main effect or 2-factor interactions, but 2-factor interactions are aliased with 3-factor interactions and main effects are aliased with 4-factor interactions.

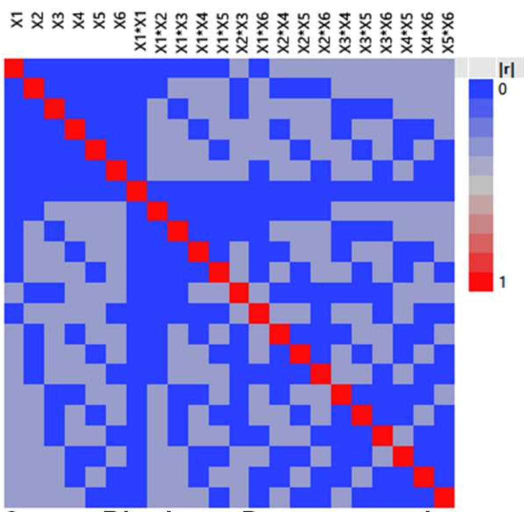
DEFINITIVE SCREENING DESIGNS

- *Definitive* – unambiguous identification of active main effects and quadratic effects, and under some conditions, 2-factor interactions
- AKA 3-level screening designs: each run includes exactly one factor at its center point, in addition to a traditional center point run
- Minimum number of runs is $2m + 1$ (m=the number of factors)
  - Recent work (Errore, et al. 2017) recommends augmentation with 4 or 8 additional runs to increase detection of 2nd order effects

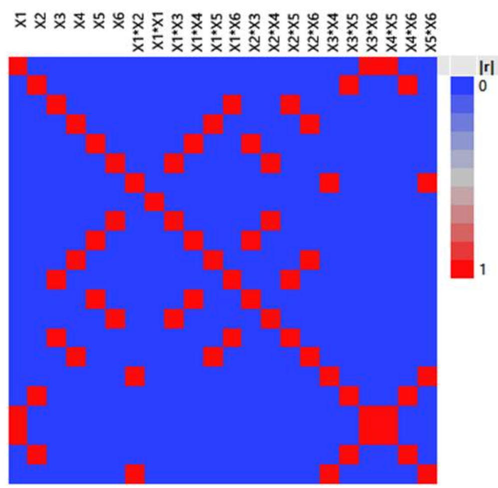
DEFINITIVE SCREENING DESIGNS

- Main effects are independently estimable.
  - In Resolution III FFs, main effects are biased by active 2-factor interactions (2FIs), whether the 2FIs are in the model or not
- No two-factor interaction is completely confounded with any other two-factor interaction. However, a 2FI might be correlated with other two-factor interactions.
- All quadratic effects are estimable and orthogonal to main effects, and not completely confounded with 2FIs
  - not simply "curvature" like in Resolution III-IV FFs with center point run(s)
- With ≥6 factors, can efficiently estimate *all possible full quadratic models involving ≤3 factors*

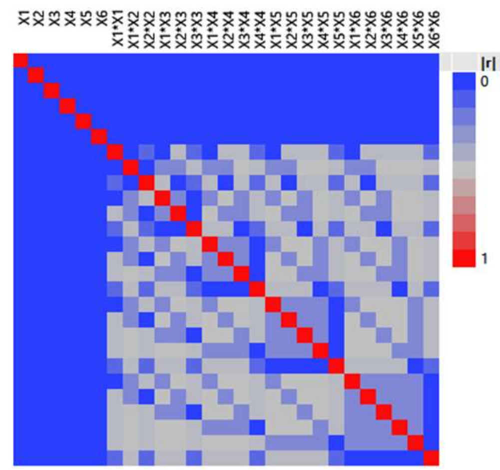
6-FACTOR DESIGNS



13-run Plackett-Burman: only main effects are independently estimable; all 2FIs are heavily confounded with MEs; cannot estimate quadratic terms – only "curvature"

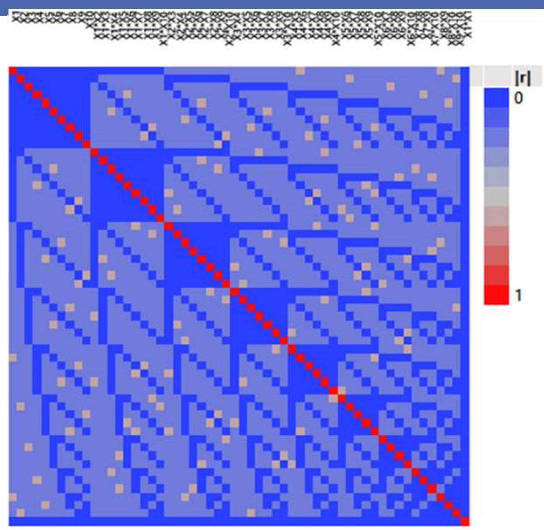


9-run Resolution III FF: MEs independent only of other MEs; cannot estimate quadratic terms – only "curvature"

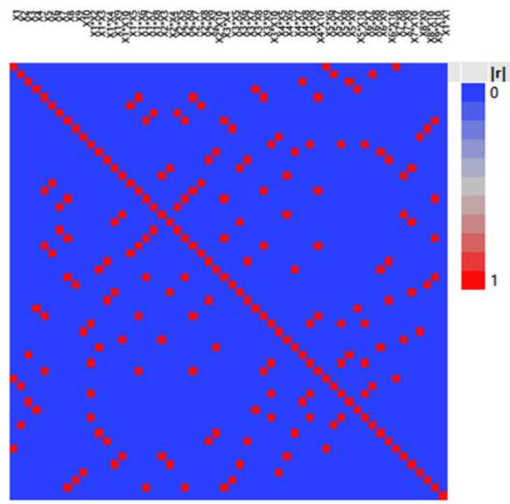


13-run DSD – all MEs + quadratic terms estimable; 2-factor interactions correlated *and* can fit RSM model if ≤3 active effects!

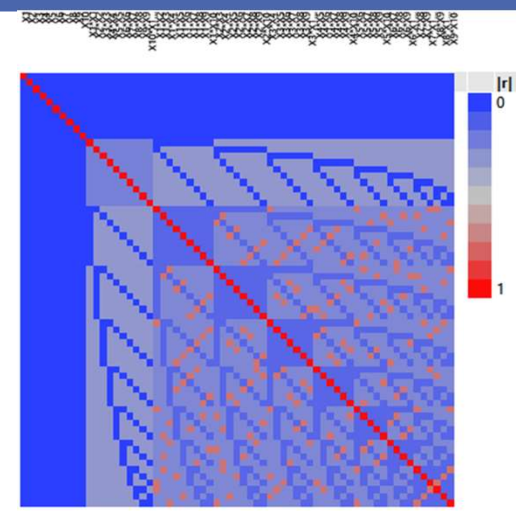
10-FACTOR DESIGNS



21-run Plackett-Burman: only main effects are independently estimable; all 2FIs are heavily confounded with MEs; cannot estimate quadratic terms – only "curvature"



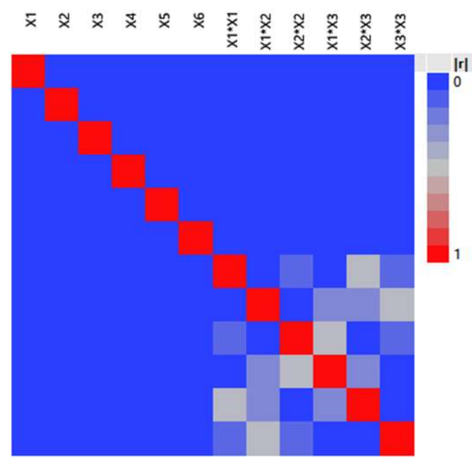
17-run Resolution III FF: MEs independent only of other MEs; cannot estimate quadratic terms – only "curvature"



21-run DSD – all MEs + quadratic terms estimable; 2-factor interactions correlated *and* can fit RSM model if ≤5 active effects!

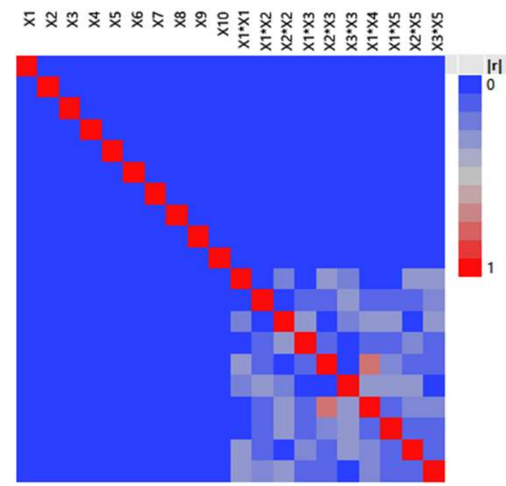
EFFICIENT 2ND ORDER MODELS WITH DSD

- If <½*m* active factors, DSDs for *m* factors can efficiently fit full 2nd order models
- Corresponding classical RSM design for 6 factors requires ≥46 runs



6-factor, 13-run DSD-> full 3-factor 2nd order model

10-factor, 21-run DSD-> full 5-factor 2nd order model



DEFINITIVE SCREENING DESIGNS

- JMP 13 has new "fit definitive screening design" feature to aid with model selection
  - Uses "Effective Model Selection for DSDs" (Jones & Nacthsheim 2016)
- Assuming effect sparsity < 0.5, DSDs are highly sensitive to detecting 2nd order effects
  - If sparsity is > 0.5, augmentation of DSD is recommended (see 2017 paper)

MODERN DOE – SUMMARY

- Efficient, cost-saving designs with same or better precision compared to traditional DOE
- Require specialized software *and* technical understanding to generate design
- DSD – be careful with model selection & analysis with many factors

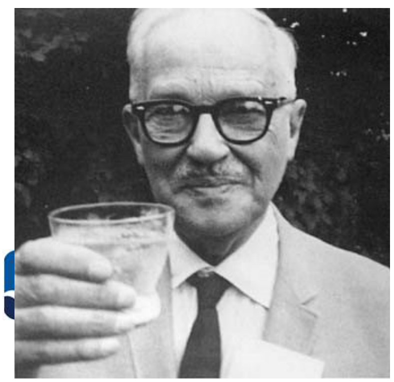
ADVANCES IN STATISTICAL ANALYSIS

BAYESIAN STATISTICS

$$
\begin{aligned}
p(Y^* \mid Y) &= \iint p(Y^* \mid \mu, \phi, Y) p(\mu, \phi \mid Y)\, d\mu\, d\phi \\
&= \iint p(Y^* \mid \mu, \phi) p(\mu, \phi \mid Y)\, d\mu\, d\phi
\end{aligned}
$$

CLASSICAL STATISTICS

….Significance testing, hypothesis testing, p-values, confidence intervals….



Probability of observed data given a parameter value

POSTERIOR PREDICTIVE DISTRIBUTION

$$p(Y^* \mid Y) = \iint p(Y^* \mid \mu, \phi, Y) p(\mu, \phi \mid Y)\, d\mu\, d\phi$$
$$= \iint p(Y^* \mid \mu, \phi) p(\mu, \phi \mid Y)\, d\mu\, d\phi$$

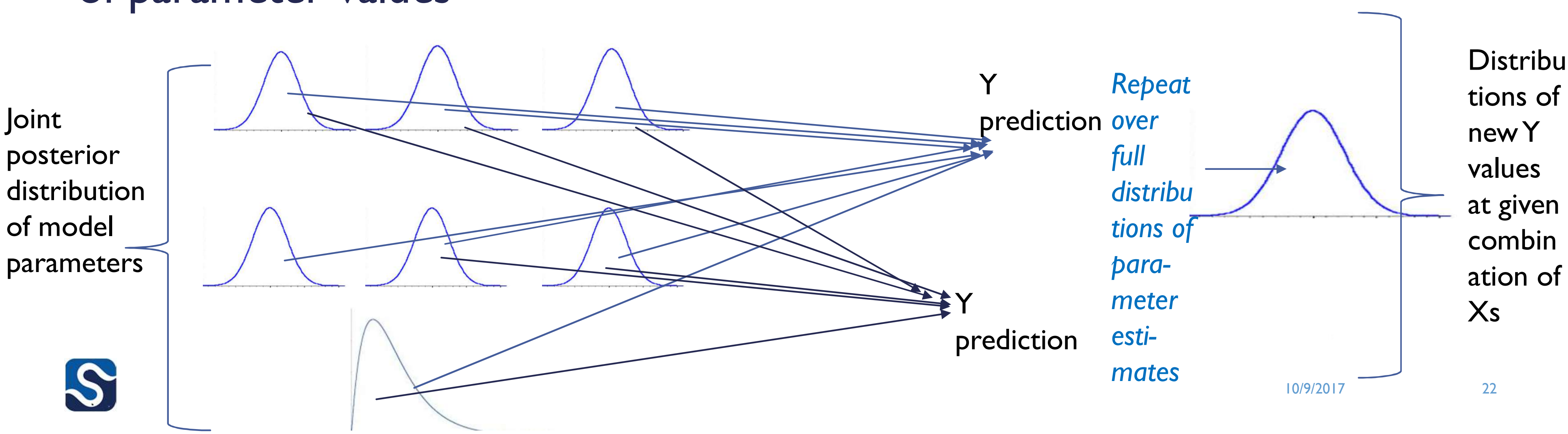
- The posterior distribution is a probability distribution that represents your updated beliefs about the parameter after having seen the data
  - Posterior distribution ∝ prior distribution × likelihood (data)
- The *posterior predictive distribution* is the distribution of *results*, given this updated belief about parameters

BAYESIAN PREDICTION

- Bayesian prediction is not the same as simulations
- Takes into account all model uncertainty + data variability
  - Explores full joint distribution of model parameters

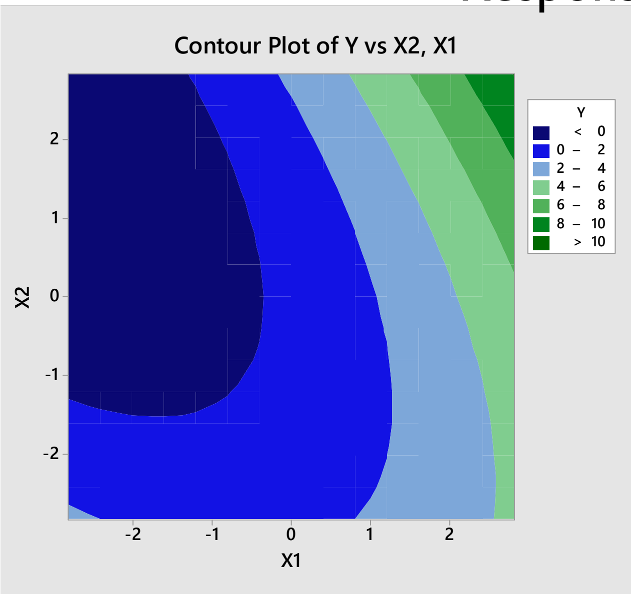
BAYESIAN PREDICTION

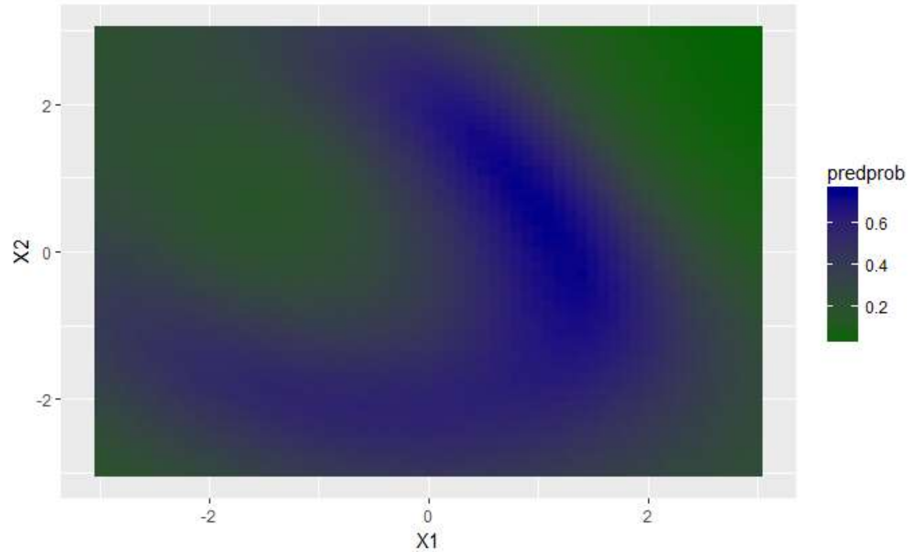
The "posterior predictive distribution" reflects true prediction of future *results*, unlike Monte Carlo simulation using point estimates of parameter values



EXAMPLE – MEAN CONTOURS PLOTS VS PROBABILITY PLOTS

Response Surface 2 factors, Y specifications 1-6





Y = 0.404 + 2.067 X1 + 0.230 X2 + 0.639 X1*X1 + 0.545 X2*X2 + 0.864 X1*X2

OTHER APPLICATIONS OF ACCURATE PREDICTION

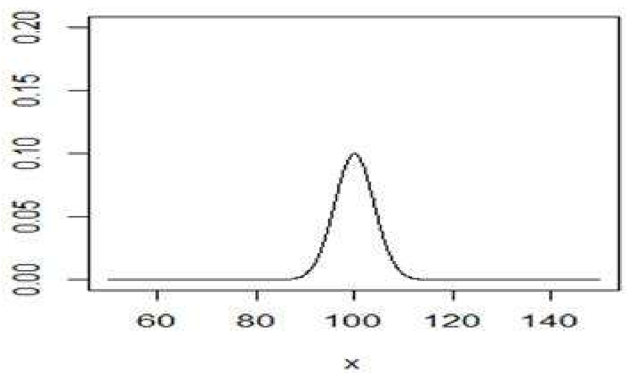
- Control limits
- Final/intermediate product specifications
- Trending (e.g. pharmaceutical stability)
- Comparability analyses

OTHER ADVANTAGES

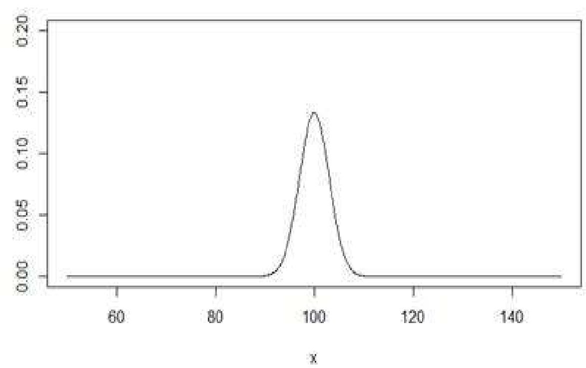
- Can model complex sampling hierarchy (multiple components, such as between, within batches, measurement error)
- Systems modeling (holistic prediction across multiple process steps; output of one unit operation becomes prior information for next)
- Use of prior information

PRIOR INFORMATION: INCREASE PRECISION WHEN DATA ARE LIMITED

Prior Distribution μ=100, s=4



Observed Distribution μ=100, s=3



4 Observed Results (sampled from normal distribution with mean=100, s = 3)

Frequentist 95/95 Tolerance interval [77.8 - 126.6]   (Frequentist prior is -∞, + ∞)
Bayesian 95/95 Tolerance Interval   [89.6 - 111.2]

more accurate prediction of performance !

BAYESIAN STATISTICS

$$p(Y^* \mid Y) = \iint p(Y^* \mid \mu, \phi, Y) p(\mu, \phi \mid Y) \, d\mu \, d\phi$$
$$= \iint p(Y^* \mid \mu, \phi) p(\mu, \phi \mid Y) \, d\mu \, d\phi$$

REFERENCES

- Jones & Nachtsheim, Journal of Quality Technology 2011
- Errore, Jones, et. al., Journal of Quality Technology 2017

BACKUP SLIDES


change
technology changes
exponentially…
…but
organizations change
logarithmically
this
challenge
grows
harder
with time
management
must choose carefully
which changes are adopted
time

OPTIMAL DESIGNS

- 1970s Optimal Design (algorithms)

  (theory is much older – 1918, Kirstine Smith)
- Allocation of experimental runs "optimal" with respect to a statistical criterion
  - Most well-known, D-Optimal designs minimize the variance-covariance matrix
- Many different optimality criteria exist, each with their own pros & cons

OPTIMAL DESIGNS

- Require **fewer runs** than non-optimal design, with same precision, for a given model
- Optimality criteria are **model-dependent** (must specify a model in advance)
- Confusion due to **lots of optimality criteria** (D-, G-, A-, I-, T-, E-, C-, V- … alphabet soup!)
- *But*, on the other hand, **universal optimality** can be shown under certain conditions (Kiefer (1975), Yeh (1986))