Dimensional Analysis & Its Applications in Statistics

Dennis Lin
Department of Statistics
The Pennsylvania State University

May, 2013

Agenda

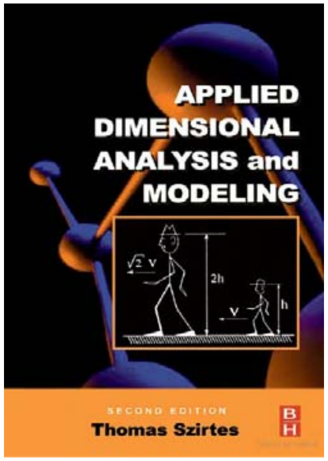
- What is Dimensional Analysis (DA)
- Illustrative Example
- Case Study: Cherry Tree (DA for Analysis)
  - Data Analysis without DA
  - Data Analysis with Dimensional Analysis
- Case Study: Paper Helicopter (DA for Design)
  - Design without DA
  - Design with Dimensional Analysis
- Lessons Learn
- Future Research Issues

Key References

Szirtes T. (1997)
"Applied Dimensional Analysis & Modeling."
Buckingham's 1914 paper

Albrecht MC, Nachtsheim CN, Albrecht TA & Cook RD (2011) "Experimental Design for Engineering Dimensional Analysis."

Davis T. (2011) "Dimensional Analysis in Experimental Design."



Pioneer Work by Buckingham

- Buckingham, E. (1914). "On physically similar systems; illustrations of the use of dimensional equations". Physical Review 4 (4): 345-376.
- Buckingham, E. (1915). "The principle of similitude". Nature 96 (2406): 396-397.
- Buckingham, E. (1915). "Model experiments and the forms of empirical equations". Transactions of the American Society of Mechanical Engineers 37: 263-296.

Dimensional Analysis: Definition

Dimensional Analysis—a tool to find relationships among physical quantities by using their dimensions.

- The dimension of a physical quantity has units.
- Quantities of different dimensions can not add, but they can multiply each other to form a derivative quantity.

Dimensional Analysis: Wikipedia

- Check the plausibility of derived equations and computations
- Form reasonable hypotheses about complex physical situations that can be tested by experiment or by more developed theories of the phenomena
- Categorize types of physical quantities and units based on their relations or dependence on other units, or their dimensions if any

Theoretical Base—Physics

- A physical law must be independent of the units used to measure the physical variables
  - Any meaningful equation (and any inequality) must have the same dimensions in the left and right sides
- Bridgeman's principle of absolute significance of relative magnitude
  - Formula should be the power-law form
- Buckingham's Π-theorem (1914)
  - Physical equations must be dimensionally homogeneous

DA: General Idea

$$Q_0 = f(Q_1, \ldots, Q_6) = \frac{Q_1 + Q_2}{Q_3} + Q_4 - Q_5 \log Q_6$$

- Q_1 and Q_2 must have the same dimension,
- Q_6 must be dimensionless, and
- Q_0, $(Q_1+Q_2)/Q_3$, Q_4 and Q_5, must have the same dimension.

DA: General Idea

$$Q_0 = f(Q_1,\ldots,Q_p) \longrightarrow \pi_0 = h(\pi_1,\ldots,\pi_{p-k})$$

- A meaningful f may have lots of constraints on itself. It can not be too arbitrary.
- Reduce dimensions from p to $p-k$,

  p is the dimension of the quantities we concern
  &
  $p-k$ is the dimension of the base quantities in the problem.

These are dimensionless variables!

Illustrative Example: Ball deformation experiment

Identify dependent and independent variables

d=f(V, ρ ,D,E, γ)

d the diameter of ball imprint	[d]=L
V the velocity of the ball	$[V]=LT^{-1}$
ρ the density of the ball	$[\rho]=ML^{-3}$
D the diameter of the ball	[D]=L
E the modulus of elasticity	$[E]=ML^{-1}T^{-2}$
γ Poisson's ratio	[γ]=1

Ball deformation experiment

- Identify a complete dimensionally independent subset

  $[V]=LT^{-1}$, $[\rho]=ML^{-3}$, [D]=L

  [d]=L, $[E]= ML^{-1}T^{-2}$, [γ]=1

- Identify the dimensionless forms of variables not in the basis set

  [d]=[D], $[E]=[V^2\rho]$, $[\gamma]=[V^0]$

Dimensional Analysis

- The potential effects on responses come from combinations of considered quantities.

$$\pi_0 = h(\pi_1,\pi_2)$$

$$\pi_0 = \frac{d}{D}$$

$$\pi_1 = \frac{E}{\rho V^2}$$

$$\pi_2 = \gamma$$

- [d]=[D], [E]=[$V^2 \rho$], [γ]=[V^0], d=f(V, ρ ,D,E, γ)

$$\pi_1 = \frac{E}{V^2 \rho}, \pi_2 = \gamma, \pi_0 = \frac{d}{D}$$

- Apply Buckinghan's Π-Theorem to get DA model

$$\pi_0 = h(Q_1, \ldots, Q_p, \pi_1, \ldots, \pi_{p-k}) = h(\pi_1, \ldots, \pi_{p-k})$$

$$\frac{d}{D} = h\left(V, \rho, D, \frac{E}{\rho V^2}, \gamma\right) = h\left(\frac{E}{\rho V^2}, \gamma\right)$$

d=f(V, ρ ,D,E, γ)

d the diameter of ball imprint	[d]=L
V the velocity of the ball	[V]=LT^{-1}
ρ the density of the ball	[ρ]=ML^{-3}
D the diameter of the ball	[D]=L
E the modulus of elasticity	[E]=$ML^{-1}T^{-2}$
γ Poisson's ratio	[γ]=1

	d	V	ρ	D	E	γ
L	1	1	−3	1	−1	0
T	0	−1	0	0	−2	0
M	0	0	1	0	1	0

	d	V	ρ	D	E	γ
L	1	1	−3	1	−1	0
T	0	−1	0	0	−2	0
M	0	0	1	0	1	0

	d	V	ρ	D	E	γ
L	1	1	−3	1	−1	0
T	0	−1	0	0	−2	0
M	0	0	1	0	1	0
π_0	1	0	0	−1	0	0
π_1	0	−2	−1	0	1	0
π_2	0	0	0	0	0	1

- [d]=[D], [E]=[$V^2 \rho$], [γ]=[V^0], d=f(V, ρ ,D,E, γ)

$$\pi_1 = \frac{E}{V^2 \rho}, \pi_2 = \gamma, \pi_0 = \frac{d}{D}$$

- Apply Buckinghan's Π-Theorem to get DA model

$$\pi_0 = h(Q_1, \ldots, Q_p, \pi_1, \ldots, \pi_{p-k}) = h(\pi_1, \ldots, \pi_{p-k})$$

$$\frac{d}{D} = h\left(V, \rho, D, \frac{E}{\rho V^2}, \gamma\right) = h\left(\frac{E}{\rho V^2}, \gamma\right)$$

So…find Ψ, such that

$$\pi_0=\Psi(\pi_1, \pi_2)$$

Instead of

Find f, such that $d=f(V, \rho, D, E, \gamma)$

Fundamental Dimensions —Tim Davis

- Length (m)
- Mass (kg)
- Time (s)
- Temperature (K)
- Electric charge (C)
- Amount of matter (mol)
- Luminous intensity (cd)

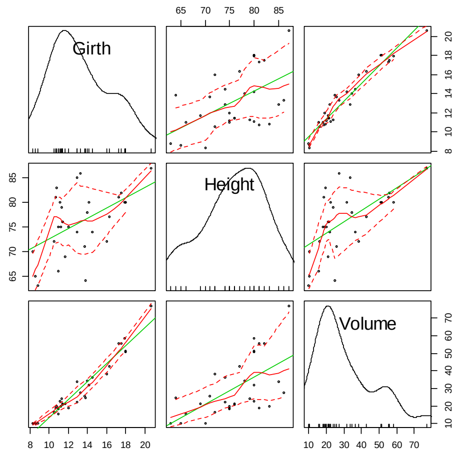
Minitab Cherry Tree Data

- 31 black cherry trees from the Allegheny National Forest
- Diameter @ 4.5 ft (inches) - d (X_1)
- Height (feet) - h (X_2)
- Marketable volume (cubic feet) - v (Y)

- Example for linear regression
  - Cook & Weisberg, 1982; Atkinson, 1985

Minitab Cherry Tree Data

	Girth(inches)	Height(feet)	Volume(feet^3)
1	8.3	70	10.3
2	8.6	65	10.3
3	8.8	63	10.2
4	10.5	72	16.4
...	...	...	...
30	18.0	80	51.0
31	20.6	87	77.0

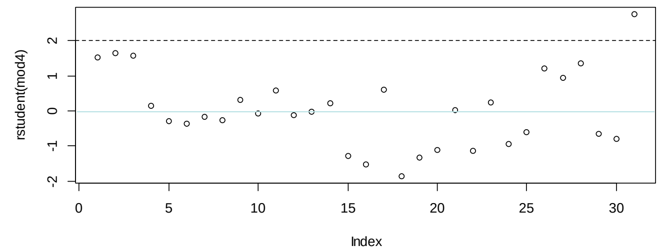


Scatter plot of the tree data

Simple linear regression

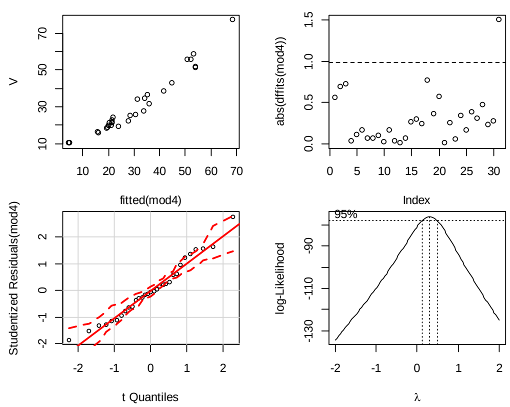
V=-58.0+56.5(3)D+0.34(0.13)H

R^2= 94%

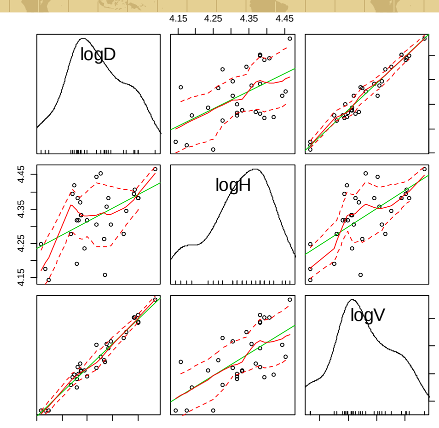


- **Studentized Residuals Plot**
- **#31 is an outlier**

Diagnostics



General fitting is good, except #31
Transformation?

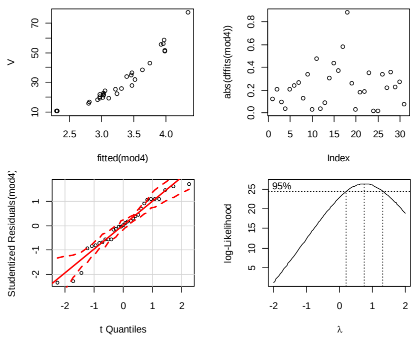


Scatter plot of the log transformed data
Better linear relationship

log transformed linear regression

$\log(V) = -1.705 + 1.98_{(.08)}\log(D) + 1.12_{(.20)}\log(H)$

$R^2 = 0.995$, and #31 is no longer an outlier



Re-Set the coefficients

$$\log(V) = C + 2\log(D) + 1\log(H)$$

$$\hat{V} = 0.3036(.004)D^2H$$

$$= 0.3866(.005)AH$$

$$A = \pi r^2 = \pi D^2/4$$

$R^2 = 99.5\ \%$

Box-Cox transformation

$$\hat{\lambda} = 0.3066 \approx 1/3$$

$$\sqrt[3]{\hat{V}} = 1.824(0.07)D + 0.014(0.001)H$$

$R^2 = 99.93\%$

- Both models are highly efficient.

Dimensional Analysis review

- Procedure:
  - Determine the inputs and their dimensions
  - Determine the base quantities
  - Transform inputs into dimensionless quantities by using base quantities
  - Re-express the estimating functions

Dimensional Analysis

Variable	Units
V	ft^3
H	ft
A~D^2	ft^2

Buckingham's Π-theorem:

relationship only include two dimensionless variables

Procedure

Get dimensionless variables

$$\Pi_V = VH^{\beta}; \Pi_A = AH^{\gamma}.$$

$$\Pi_V = \frac{V}{H^3}; \Pi_A = \frac{A}{H^2}.$$

Estimate functions

$$\Pi_V = f(\Pi_A, H) = f(\Pi_A)$$

$$\Pi_V = k(\Pi_A)^{\delta}$$
$$\Leftrightarrow V = kA^{\delta}H^{3-2\delta}$$
$$\Leftrightarrow \log V = C + \delta \log A + (3-2\delta)\log H$$

$$\Pi_V = (a\sqrt{\Pi_A} + b)^3$$
$$\Leftrightarrow \sqrt[3]{V} = aD + bH$$

Special Case-I

$\Pi_V = k(\Pi_A)^{\delta}$

Set δ=1

$$\Pi_V = 0.3850(.005)\Pi_A$$
$$V = 0.3850(.005)AH$$

- This is the same as the log transformation

$$\hat{V} = 0.3866(.005)AH$$

- R^2=99.5%

Special Case-II

$\Pi_V = k(\Pi_A)^{\delta}$

Set δ=3 and k=1

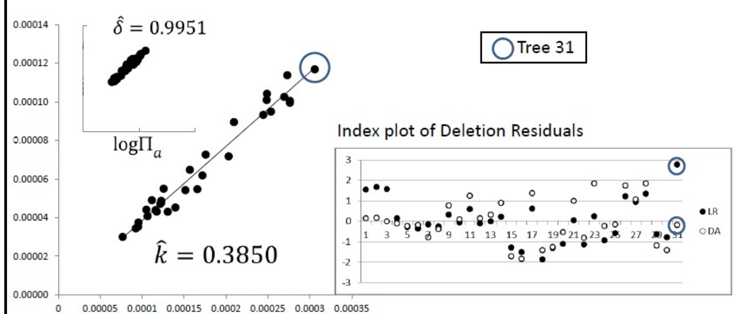
$$\Pi_V = (\alpha\Pi_A^{1/2} + \beta)^3$$

- This is the same as the linear model

$$\sqrt[3]{\hat{V}} = 1.824(0.07)D + 0.014(0.001)H$$

- R^2=99.93%

Diagnosis



Summary on cherry tree

- Lesson learned:
  - Reduce input variables from 2 to 1
  - No lose on any information
  - cover traditional models
  - Similar results
- Comments:
  - No harms to incorporate DA before analysis
  - Better interpretation

Related Issue:

Error Structure

Model Fitting & Diagnosis

Error Structure

Assume model: $\pi_0 = \prod \pi_i^{\beta_i} \cdot \varepsilon$

i.e., $(\log \pi_0) = \sum \beta_i (\log \pi_i) + \log \varepsilon$

We have $E(\widehat{\log} \pi_0) = \log \pi_0$

However, $E(e^{\widehat{\log} \pi_0}) \neq e^{E(\widehat{\log} \pi_0)}$

i.e., $E(\hat{\pi}_0) \neq \pi_0$

Statistical Inference: DA Model

$$y = \beta_0 x_1^{\beta_1} x_2^{\beta_2} \cdots x_p^{\beta_p} \varepsilon$$

$$\min_{\alpha,\beta_1,\beta_2} \sum_i (\log y_i - \log \alpha - \beta_1 \log x_{1i} - \beta_2 \log x_{2i})^2$$

$$\min_{\alpha,\beta_1,\beta_2} \sum_i |\log y_i - \log \alpha - \beta_1 \log x_{1i} - \beta_2 \log x_{2i}|$$

$$\min_{\alpha,\beta_1,\beta_2} \sum_i \left(\left| \frac{y_i - \alpha x_{1i}^{\beta_1} x_{2i}^{\beta_2}}{\alpha x_{1i}^{\beta_1} x_{2i}^{\beta_2}} \right| + \left| \frac{y_i - \alpha x_{1i}^{\beta_1} x_{2i}^{\beta_2}}{y_i} \right| \right)$$

$$\min_{\alpha,\beta_1,\beta_2} \sum_i \left(\left| \frac{y_i - \alpha x_{1i}^{\beta_1} x_{2i}^{\beta_2}}{\alpha x_{1i}^{\beta_1} x_{2i}^{\beta_2}} \right| \times \left| \frac{y_i - \alpha x_{1i}^{\beta_1} x_{2i}^{\beta_2}}{y_i} \right| \right)$$

Statistical Inference

Model: $y = \beta_0 x_1^{\beta_1} x_2^{\beta_2} \cdots x_p^{\beta_p} \varepsilon$

I. min $\left[\log\left(\frac{y}{\beta_0 x_1^{\beta_1} x_2^{\beta_2} \cdots x_p^{\beta_p}} \right) \right]^2$

II. min $\left[\frac{y}{\beta_0 x_1^{\beta_1} x_2^{\beta_2} \cdots x_p^{\beta_p}} - 1 \right]^2$

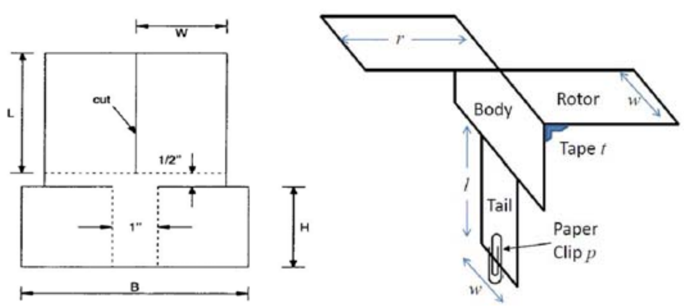
III. min $\left[\frac{\beta_0 x_1^{\beta_1} x_2^{\beta_2} \cdots x_p^{\beta_p}}{y} - 1 \right]^2$

Paper Helicopter:

Dimensional Analysis for Design of Experiment

What is a paper helicopter?

Goal: maximize the landing time

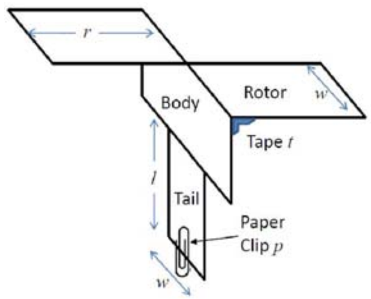


Paper Helicopter

- Literature Review
  - Johnson et al (QE 2006)
  - Box & Liu (JQT 1999)
    - 1st experiment
    - 2nd experiment
  - Annis (AS 2005)
- Dimensional Analysis on Paper Helicopter
  - Tim Davis (2011)

Johnson (QE 2006)

- Input: Two levels
  1. Paper type
  2. Body length
  3. Body width
  4. Wing length
  5. Paper clip
  6. Body tape
  7. Joint tape
- Output: Flight time



Johnson (QE 2006)

Design: (Seven two-level input variables)

- Half Fractional Factorial (2^{7-1} design)
- Two replicates (total of $2^{7-1}*2=128$ runs)
- Resolution VII: all main, two-factor, and three-factor interaction effects are clear.

main ~ six-way;
two-way ~ five-way;
three-way ~ four-way.

Johnson (QE 2006)

Term	Effect
Constant	
Paper	−0.14734
Clip	−0.12797
Bodytape	−0.05828
Width	−0.17797
Length	−0.16391
Wing	0.49297
Paper*Clip	−0.04484
Paper*Width	−0.05172
Clip*Length	−0.04984
Length*Wing	−0.05516

Significant factors:

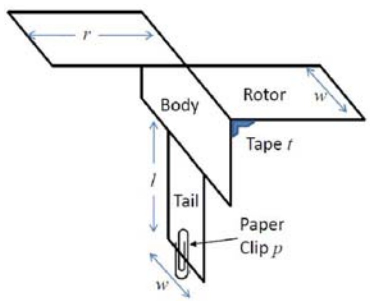
Large main effect;
Moderate two-way effect;
NO higher order effect.

Johnson(QE 2006): Conclusion

- Case study of "Six Sigma" Black Belt project
- Build best helicopters (air force)
- Consider many variables (7 and interactions)
- Typical routine to do design and analysis
- Step by step reasoning to maximize
- Limited budget

Box & Liu (JQT 1999)

- Input: Two-level
  1. Paper type
  2. Wing length
  3. Body length
  4. Body width
  5. Fold
  6. Taped body
  7. Clip
  8. Taped wing
- Output: Flight time



Box & Liu (JQT 1999)

Design: (8 two-level input variables)

- Fractional 2-level Resolution IV (a 2_{IV}^{8-4} design)
- 4 replicates (a total of 4*16=64 runs)

- Wing length (3 inches vs. 4.75 inches)
- Body length (3 inches vs. 4.75 inches)
- Body width (1.25 inches vs. 2 inches)

Box & Liu (JQT 1999)

Significant Effects (No interactions)

Variables	Mean time	Dispersion
Paper type	+	+
Wing length (l)	+	-
Body length (L)	-	+
Body width (W)	-	+
Fold	+	+
Taped body	+	+
Paper clip	-	-
Taped wing	-	+

Box & Liu (JQT 1999)

- Resulting Model:

  $$\hat{y} = 223 + 28l - 13L - 8W$$

  y in centiseconds

- Further optimization:
  - Linear assumption: coefficients change according to specific *l, L, W.*
  - Search the maximum by experiments according to steepest ascent.

Box & Liu (JQT 1999)

- Series designs for searching optimum point
- Not "one-shot" but "sequential learning"
- Steepest Ascent
- Optimum means longest flight time with minimum variance
- Higher order designs and final optimum of 416 cent-sec.

Annis (AS 2005)

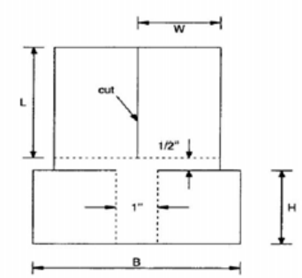
- Input:
  - Base length B
  - Base height h
  - Wing length L
  - Wing width W
- Model: (Physics)

$$E(Y) = \beta_0 + \beta_1 \log\left(\frac{\beta_2^{\ 2}}{LW} + LW\right)$$

$$E(Y) = E[\log(t)] + \frac{1}{2}\log(S)$$

$$S = BH + (2L+1)W$$

- Output: Flight time



Annis (AS 2005)

- Design:
  - Three-level full factorial design for L and W (3^2 design)
  - D=15.5 feet
  - Response surface
- Result:
  - Get 4.34 seconds when L=6 W=1.81. (Theoretically based on response surface)

Annis (AS 2005)

- Incorporate physical derivation before design
- Engineers provide theory for guidance
  - Parts we believe; Parts we doubt
- Statisticians provide data for validation
  - Parameter estimation; Question physics
- Better than full factorial design
- Extrapolation
- Nonlinear response and drop lower order terms

Literature Review: without DA

	Johnson(QE 2006)	Box & Liu(JQT 1999)	Annis(AS 2005)
Input	+Paper type(-) +Taped body(-) +Taped wing(-) +Clip(-) +Interaction	+Paper type(-) +Taped body(-) +Taped wing(-) +Clip(-) +Fold(-) (+Wing area & ratio)	Body Length(-) Body Width(-) Wing length(+) +Wing Width(dip)
Design	2-level(-1,+1) Half factorial (VII)	2-level Fractional (IV) & full factorial	3-level Full factorial
Number of Runs	64*2	16*4&16	9
Final Model	Y=2.11-0.089W-0.082L+0.246w	Y=223+28l-13L-8W Y=326+8A-17L	Y=6.147-.79log(358/lw+lw)-.5log(LW+(2l+1)w)
Optimum value	2.847s 8'=2.44m	4.16s 8'6"=2.59m	4.34s 15'6"=4.72m
Lessons	Interactions	Sequential learning	Physical insight
Key variables	Body length(-)	Body width(-)	Wing length(+)

Paper Helicopter: with DA

- Input $T = F_1(m, g, r, c_d, \rho, h)$
  - *m*: Mass, *g*: Gravity const., *r*: Wing Length,
  - C_d: Viscosity const., ρ: Density, *h*: Height
- Prior reduction $T = \frac{h}{v}, c_d$

  $v = F_2(m, g, \rho, r)$
- DA $\Phi_v = \frac{h}{T\sqrt{gr}}; \Psi_m = \frac{m}{\rho r^3}$

  $\Phi_v = F_3(\Psi_m)$

Paper Helicopter: DA

- Model: $\Phi_v = F_3(\Psi_m)$
- Design:
  - 4 levels,
  - 3 replicates,
  - equal separation
- Result:

$$T = \frac{hr}{0.859}\sqrt{\frac{\rho}{mg}} \qquad m = (3.09)g \quad \rho = 1204 g/m^3$$

$$\Phi_v = 0.859\sqrt{\Psi_m} \qquad r = (0.14)m \quad g = 9.8 N/kg$$

$$T = (5.18)s \quad h = 5.3m$$

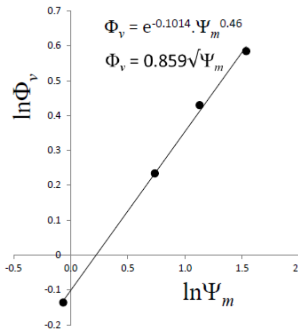
Paper Helicopter: DA

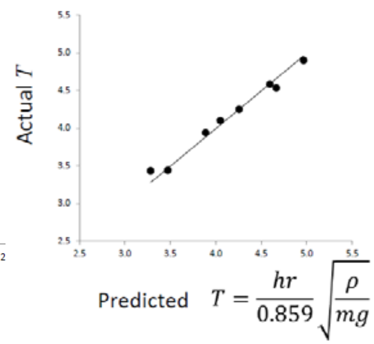
#	Paper (gsm)	'copter mass (m)	Rotor radius* (r)	Ψ_m $m/(\rho r^3)$	Flight Time (T)**	Φ_v $h/(T\sqrt{gr})$
❶	80	3.09g	140mm	**0.937**	5.18s	**0.873**
❷	120	4.34g	120mm	**2.087**	3.87s	**1.264**
❸	100	3.72g	100mm	**3.088**	3.48s	**1.537**
❹	160	5.59g	100mm	**4.642**	2.98s	**1.795**

* Rotor width fixed at 52.5mm

** Average of 3 flights recorded twice

Paper Helicopter: DA





$$T = \frac{hr}{0.859}\sqrt{\frac{\rho}{mg}}$$

	Previous (Without DA)	Davis (With DA)
Variables	Two or three levels	Continuous (interpolate and extrapolate)
Design	On variables	On dimensionless transformations 4 or 5 levels
Result	Wing length(+) Body length(-) Body width(-) Area(+) Ratio(?)	Wing length(+) Body length(-) Body width(-) Area(+) Ratio(?)
Optimum	4.34s 4.7m v=1.09m/s	5.97s 5.3m v=0.89m/s
Opt. Point	l=15.2cm w=4.60cm m=A4 sheet	l=14cm w=7cm m=3.09g
Estimate function	Y=6.147-.79log (358/lw+lw)-.5log(LW+(2l+1)w)	$Y = \frac{h}{0.6016}\sqrt{\frac{\rho lw}{mg}}$
Compared model	Full factorial	Confirmation runs

Summary on paper helicopter

- Lessons learned: from design
  - Reduce input from 4 to 1, and 5 to 2
  - Save costs if base designs on transformed dimensionless variables (separate covariate space)
  - Similar results
- Comments:
  - Save costs even small reductions
  - Group variables
  - Scalability

Related Issue:

There are many combination of Q_i's to provide the same value of Π_j,
which combination is Optimal?

General Comments

- Engineers provide theory for guidance.
  - Use physical prior knowledge
  - Only test the parts with unknown physical structure
- Statisticians provide data for validation.
  - Check the validity of physical assumption
  - Recommend further experiments

(Annis 2006)

Pros and Cons

- Pros:
  - Nature of relationships (Not always linear)
  - Priori reduction
  - Scalability
- Cons:
  - Physical knowledge
  - Possible severe problems if important related variables were missing

Agenda

- What is Dimensional Analysis (DA)
- Illustrative Example
- Case Study: Cherry Tree (DA for Analysis)
  - Data Analysis without DA
  - Data Analysis with Dimensional Analysis
- Case Study: Paper Helicopter (DA for Design)
  - Design without DA
  - Design with Dimensional Analysis
- Lessons Learn
- Future Research Issues

Join Us!

There are whole lots more to be done!

Send $500 to

- Dennis Lin
  University Distinguished Professor
  317 Thomas Building
  Department of Statistics
  Penn State University

- +1 814 865-0377 (phone)
- +1 814 863-7114 (fax)
- DennisLin@psu.edu

(Customer Satisfaction or your money back!)

Error Structure

Assume model: $\pi_0 = \Pi \pi_i^{\beta_i} \cdot \varepsilon$

i.e., $(\log \pi_0) = \sum \beta_i (\log \pi_i) + \log \varepsilon$

We have $E(\hat{\log} \pi_0) = \log \pi_0$

However, $E(e^{\hat{\log} \pi_0}) \neq e^{E(\hat{\log} \pi_0)}$

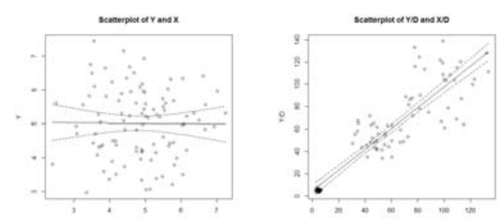
i.e., $E(\hat{\pi}_0) \neq \pi_0$

Dependence: Before & After

- Y & X are independent
  ➔ Y|D & X|D are independent
- Y & X are dependent
  ➔ Y|D & X|D are dependent
- Y & X are independent
  ➔ Y|D & X|D are dependent
- Y & X are dependent
  ➔ Y|D & X|D are independent

Dependence Before/After DA

- Before, Y and X uncorrelated
- After, Y/D and X/D correlated
- Spurious correlation. Conversed Result.



Dimension: Variable & Constant

- Physical constants have dimensions.
- Boltz-mann constant (k), gravitational constant (G), speed of light (c).
- No variations. To be estimated.
- Should be included to avoid ruling out important variables during DA.

- Parameter: Stat vs Physics

Missing Key Variables

- Missing key variables in DA
  -->associated deletion of others
- Critical but not fatal
- Worst Scenario: one per basis quantity
- If basis quantity d is only contained by Q, cautious of missing quantities.

Scalability

- Scalable because power law form: $\frac{Q_1^2}{Q_2}$
- Rarely available in other models
- Still need to check extrapolation:
  - Basis quantities usually scale
  - Some quantities (constants) do not scale
  - After DA, some lie out of design space

Quantity Property

- Power law --> 0 in the denominator ?
- Physical quantity can be 0 or very small.

- Continuous quantity.

But could well be for

- Ordered quantity.
- Categorical quantity.

DA vs PCA

- Dimension Reduction Technique
- DA based on physical law
- PCA based on data

- Robustness in missing key variables?

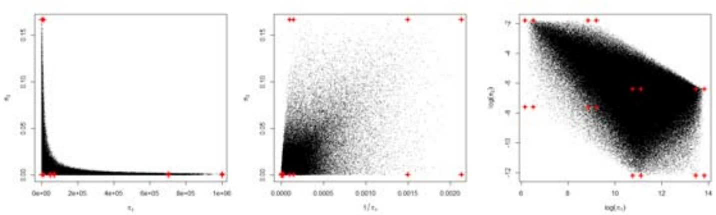
How will Bayesian do here???

- DA: Physical Prior on Coefficients
- Bayesian: Prior needed for Coefficients

- Treat DA from Bayesian point of view
- Take physical knowledge as Bayesian prior

Irregular Design Support

- Two types of support for DA variables:
- Hyperbolic ; Regular
- Log-transformation



Choices of Basis Quantities

- Basis Quantities – Subset of Variables
- Not Unique! Different Result?
- Optimal Choice; Optimal Criterion

- Canonical Choice:
  - Conventions
  - Scale of Systems

Multivariate Control Chart (X_1, X_2)

- Two Individual Control Charts
  - for both $\mathbf{X}_1$ and $\mathbf{X}_2$.
- One Multivariate Chart
  - Hotelling $\mathbf{T}^2$ Chart
- One DA Control Chart
  - Control Chart on $\frac{X_1}{X_2^\alpha}$ (eg, BMI=$\frac{Kg}{m^2}$)

Initial Simulation Setup

- Individual Chart

  $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$
- Multivariate Chart

  $$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_{12} \\ \rho\sigma_{12} & \sigma_2^2 \end{bmatrix}\right)$$
- DA Chart $\frac{X_1}{X_2}$: Ratio of two normals

Ratio of two Normals *(Cedilnik et al., 2004)*

Theorem 2. *The probability density for* $Z = X/Y$, *where* $[X\ \ Y]^T : N(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho \neq \pm 1)$ *is expressed as a product of two terms:*

$$p_Z(z) = \frac{\sigma_X \sigma_Y \sqrt{1-\rho^2}}{\pi(\sigma_Y^2 z^2 - 2\rho\sigma_X\sigma_Y z + \sigma_X^2)} \cdot \left[\exp\left(-\frac{1}{2}\cdot \sup R^2\right)\cdot\left(1 + \frac{R\cdot\Phi(R)}{\varphi(R)}\right)\right] =$$

$$= \frac{\sigma_X \sigma_Y \sqrt{1-\rho^2}}{\pi(\sigma_Y^2 z^2 - 2\rho\sigma_X\sigma_Y z + \sigma_X^2)} \cdot \left[\exp\left(-\frac{1}{2}\cdot \sup R^2\right) + \sqrt{2\pi}\cdot R\cdot\Phi(R)\cdot\exp\left(-\frac{1}{2}\cdot\left[\sup R^2 - R^2\right]\right)\right] \quad (2.2)$$

X_1 and X_2 are dependent
(ρ>0 and $X_1 \leq X_2$)

	Preferable Chart
x_1↗ x_2→	x_1-bar & x_2-bar
x_1↘ x_2→	x_1-bar & x_2-bar
x_1→ x_2↗	T^2 Chart
x_1→ x_2↘	DA Chart
x_1↗ x_2↗	T^2 Chart
x_1↘ x_2↘	DA Chart
x_1↘ x_2↗	DA Chart
x_1↗ x_2↘	DA Chart

X_1 and X_2 are independent
(ρ=0 and $X_1 \le X_2$)

	Preferable Chart
$x_1\nearrow$ $x_2\rightarrow$	x_1-bar & x_2-bar
$x_1\searrow$ $x_2\rightarrow$	x_1-bar & x_2-bar
$x_1\rightarrow$ $x_2\nearrow$	x_1-bar & x_2-bar
$x_1\rightarrow$ $x_2\searrow$	x_1-bar & x_2-bar
$x_1\nearrow$ $x_2\nearrow$	T^2 Chart
$x_1\searrow$ $x_2\searrow$	T^2 Chart
$x_1\searrow$ $x_2\nearrow$	DA Chart
$x_1\nearrow$ $x_2\searrow$	DA Chart

Multivariate Control Chart *(weight, height)*

- Two Individual Control Charts
  - for both *weight* and *height*.
- One Multivariate Chart (*weight* & *height*)
  - Hotelling **T**² Chart
- One DA Control Chart
  - Control Chart on BMI= $\frac{weight}{height^2} = \frac{Kg}{m^2}$

Performance Comparison for Three Different Charts

	Preferable Chart
weight$\nearrow$ *height*$\rightarrow$	DA Chart
weight$\searrow$ *height*$\rightarrow$	T^2 Chart
weight$\rightarrow$ *height*$\nearrow$	T^2 Chart
weight$\rightarrow$ *height*$\searrow$	DA Chart
weight$\nearrow$ *height*$\nearrow$	T^2 Chart
weight$\searrow$ *height*$\searrow$	T^2 Chart
weight$\searrow$ *height*$\nearrow$	DA Chart
weight$\nearrow$ *height*$\searrow$	DA Chart

Based on BMI analysis, the conclusion is…

I am too short!