

ESTUDO DE CASO: ANÁLISE ESTATÍSTICA DE DADOS HISTÓRICOS DE PLANADORES

Kirk Lara, Prof^o Dr. Sérgio Butkewitsch Choze

Universidade do Vale do Paraíba/ Faculdade de Engenharia Arquitetura e Urbanismo, kirk.lara@gmail.com

Resumo - Este trabalho trata da análise e tratamento de dados históricos tabulados de planadores de diversos fabricantes que utilizam filosofia similar em sua construção e aspecto geométrico. Busca-se com este tratamento, encontrar os modelos matemáticos que correlacionam as variáveis envergadura, alongamento, peso e $L/D_{\text{máx}}$ apresentados através do tratamento estatístico com a aplicação da análise da regressão. De forma proposital e didática, esta análise será realizada utilizando variáveis que embora sejam declaradas como independentes, na verdade apresentam algum tipo de dependência. Pretende-se com isso avaliar o efeito de se violar uma premissa que é a condição principal para o sucesso da análise da regressão, que é a utilização de variáveis independentes.

Palavras-chave: Análise da regressão. Inferência. Estatística. Correlação. Planador.

Área do Conhecimento: Estatística Aplicada.

Introdução

A determinação de um modelo que melhor relacione dados numéricos pode ser considerada de grande importância para diversos trabalhos em diversas áreas da engenharia.

Neste trabalho trataremos de um conjunto de dados históricos de planadores de fabricantes diferentes que utilizam filosofia de construção e geometria similares em suas aeronaves.

O tratamento destes dados será realizado utilizando a técnica estatística da análise da regressão. A análise da regressão é uma técnica utilizada para modelagem e análise de dados numéricos onde se busca entender a correlação entre as variáveis dependentes e independentes do processo.

Ângulo e razão de planeio

Para entender o vôo de um planador, vamos primeiramente montar um cenário simplificado. Vamos adotar que o vôo é realizado em ar calmo, linha reta e a velocidade constante. Nestas condições surge uma trajetória de vôo com uma inclinação φ , chamada de ângulo de planeio.

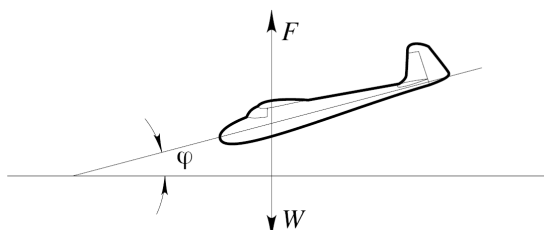


Figura 1 - Ângulo de planeio

As forças atuantes no planador são o peso W e a força aerodinâmica F .

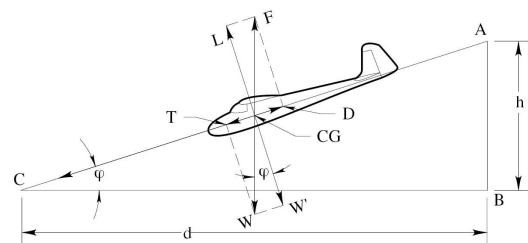


Figura 2 - Componentes do vôo

Analisando a Figura 2, vamos considerar as componentes das forças F e W em relação a duas direções, uma vertical e outra paralela à direção de vôo. As componentes de F são a sustentação L e o arrasto D . As componentes do peso W são W' e a tração T , e eles são opostos a L e D . A tração T determina o movimento ao longo da trajetória e depende do ângulo de planeio φ e do peso W . Podemos dizer então que:

$$\frac{L}{D} = \frac{d}{h} \quad (1)$$

sabendo que

$$\frac{L}{D} = \frac{C_L}{C_D} = E \quad (2)$$

e

$$\frac{d}{h} = E \quad (3)$$

A razão d/h é conhecida como razão de planeio, e o seu valor representa a eficiência aerodinâmica E . De forma recíproca, h/d representa a inclinação da trajetória p .

$$p = \frac{h}{d} = \frac{1}{E} \quad (4)$$

E máximo também é conhecido como L/D_{\max} .

Alongamento

O alongamento é representado pela equação (5).

$$AR = \frac{b}{c_m} \quad (5)$$

onde:

AR= Alongamento
b = Envergadura da asa
 c_m = Corda média

Multiplicando os termos da equação (5) pela envergadura da asa, temos:

$$AR = \frac{b^2}{S} \quad (6)$$

Durante a condição normal de vôo, a sustentação depende da distribuição de pressão na asa, onde a pressão se faz maior no intradorso e menor no extradorso. Por isso, as partículas de ar terão a tendência de escapar pelas pontas das asas das zonas de alta pressão no intradorso para as zonas de baixa pressão no extradorso, girando ao longo das pontas da asa. Este fenômeno pode ser visualizado na Figura 3.

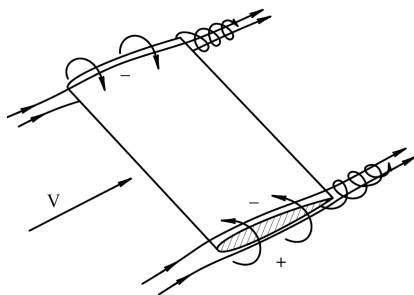


Figura 3 - Geração do arrasto induzido

Como o ar flui na direção V, as partículas de ar nas pontas das asas executarão um movimento resultante em espiral conhecido por vórtice de ponta de asa. Este vórtice é responsável por produzir um aumento no arrasto e diminuição na sustentação. Esta parcela de arrasto criada é conhecida por arrasto induzido.

O aumento no alongamento causa uma redução na corda da asa, e conseqüente redução do arrasto, que depende de dois fatores: o arrasto de perfil (C_{dp}) e arrasto induzido (C_{di}).

$$C_d = C_{dp} + C_{di} \quad (7)$$

O coeficiente de arrasto induzido é dado por:

$$C_{di} = \frac{2 \cdot (C_L)^2}{\pi} \cdot \frac{1}{AR} \quad (8)$$

O arrasto induzido é de fato, o arrasto produzido pelos vórtices das pontas da asa.

Portanto, o alongamento influencia o arrasto induzido enquanto que o arrasto de perfil permanece o mesmo.

Efeito do peso no planeio

É importante notar que o peso W não participa da equação (2). Portanto, o peso não é um fator para determinar o quanto um planador pode planar.

Quando se busca determinar o planeio para percorrer a maior distância, é necessário operar no C_L ótimo, que produz a máxima razão L/D.

No entanto, existe uma relação direta entre o peso e a velocidade de planeio. Quanto mais pesada for a aeronave, mais rápido ela terá que voar para percorrer a mesma distância.

Estudo de caso

A teoria descrita nas seções anteriores da natureza comportamental das variáveis envolvidas e suas dependências são importantes para o entendimento dos possíveis modelos e resultados obtidos com a análise estatística.

As variáveis independentes em estudo serão:

- x_1 : Envergadura (*Wingspan*)
- x_2 : Alongamento (*Aspect Ratio*)

Busca-se conhecer a lei de formação que descreve a colaboração destas variáveis para a obtenção das variáveis dependentes y_1 e y_2 , bem como a correlação entre x_1 e x_2 .

- y_1 : Peso bruto (*Gross Weight*)
- y_2 : L/D_{\max}

Observação dos dados

A tabela utilizada possui 72 amostras, com as informações de x_1 , x_2 , y_1 e y_2 completas.

Através de uma simples inspeção visual dos dados percebe-se uma aparente correlação positiva entre os dados da envergadura x_1 com o peso bruto y_1 e do alongamento x_2 com o L/D_{\max} y_2 .

Este resultado é de certa forma esperado, pois é sabido que com o aumento da envergadura, espera-se um aumento do peso bruto e com o aumento do alongamento, espera-se o aumento de L/D_{\max} .

Transformação dos dados, extração dos parâmetros estatísticos e distribuição padrão dos parâmetros

Devido ao comportamento não linear dos elementos envolvidos, utilizaremos o Modelo Quadrático Completo.

Por estarmos tratando de duas variáveis independentes e duas dependentes, o modelo de regressão linear múltipla quadrático completo em duas variáveis representativo do fenômeno pode ser representado matematicamente da seguinte forma:

$$f(x_1, x_2) = c_0 + c_1x_1^2 + c_2x_2^2 + c_3x_1 + c_4x_2 + c_5x_1x_2 \quad (9)$$

onde c_0, c_1, c_2, c_3, c_4 e c_5 são os coeficientes de regressão.

Para encontrar os valores dos coeficientes de regressão, será implementada a solução matricial definida pela equação (10).

$$\{c\} = (\{X^T\} \cdot \{X\})^{-1} \cdot \{X^T\} \{y\} \quad (10)$$

O primeiro passo para a resolução da equação (10), é determinar a matriz X (72,6).

A	TM	x_1^2	x_2^2	x_1	x_2	x_1x_2
1	1	225,00	331,24	15,00	18,20	273
2	1	306,25	292,41	17,50	17,10	299,25
3	1	306,25	295,84	17,50	17,20	301
4	1	225,00	384,16	15,00	19,60	294
5	1	225,00	453,69	15,00	21,30	319,5
6	1	400,00	595,36	20,00	24,40	488
⋮	⋮	⋮	⋮	⋮	⋮	⋮
72	1	400,00	521,67	20,00	22,84	456,8

onde

A = Amostra

TM = Termo Médio

Tabela 1 - Determinação da matriz X

Com a matriz X definida e aplicada a operação prevista na equação (10) obtém-se a matriz de variância e covariância.

1,2E+01	2,4E-02	2,6E-03	-9,8E-01	-1,9E-01	1,3E-03
2,4E-02	5,4E-04	1,3E-04	-7,3E-03	3,4E-03	-5,3E-04
2,6E-03	1,3E-04	7,7E-05	-3,4E-04	3,1E-05	-2,0E-04
-9,8E-01	-7,3E-03	-3,4E-04	1,8E-01	-5,6E-02	3,9E-03
-1,9E-01	3,4E-03	3,1E-05	-5,6E-02	5,3E-02	-2,8E-03
1,3E-03	-5,3E-04	-2,0E-04	3,9E-03	-2,8E-03	6,6E-04

Tabela 2 - Matriz de variância e covariância

Os resultados da análise da regressão são apresentados nas próximas tabelas.

Variáveis	Significância (1-p)	Coefficiente de Regressão
x_1^2	99,28%	4,142321
x_2^2	98,65%	1,413169
x_1	58,27%	-6,34785
x_2	75,32%	11,30299
$x_1 \cdot x_2$	99,31%	-4,61791

Tabela 3 - Resultado da análise da regressão para o peso bruto (y_1)

Variáveis	Significância (1-p)	Coefficiente de Regressão
x_1^2	83,97%	0,071947921
x_2^2	91,33%	0,037479338
x_1	62,05%	0,406473491
x_2	97,08%	1,377984191
$x_1 \cdot x_2$	93,06%	-0,119186278

Tabela 4 - Resultado da análise da regressão para o L/D_{max} (y_2)

Os valores em negrito na Tabela 3 e Tabela 4 referem-se aos valores que excederam 90,00%.

Teste de hipóteses

Para determinar o quanto o modelo encontrado representa o fenômeno em análise, faremos o teste de hipótese utilizando a análise da variância.

Prob(F) y_1	Prob(F) y_2
4,37E-17	3,58E-23

Tabela 5 - Prob(f) para y_1 e y_2

Com os valores encontrados na Tabela 5 sendo igualados a zero, conclui-se que todas as variáveis tem significativa participação no modelo. Este fato se comprova ao analisar a Tabela 3 e Tabela 4, onde todas as variáveis apresentam participação no processo.

$R^2 y_1$	$R^2 y_2$
0,721737	0,81902

Tabela 6 – R^2 para y_1 e y_2

A Tabela 6 apresenta os valores de R^2 para cada modelo encontrado.

Conclusão da análise da regressão

Como um dos resultados da análise da regressão, extraímos os modelos que correlacionam as variáveis dependentes às suas variáveis independentes.

$$\hat{y}_1 = 207,35 + 4,14x_1^2 + 1,41x_2^2 - 6,35x_1 + 1,130x_2 - 4,62x_1x_2 \quad (11)$$

$$\hat{y}_2 = 9,30 + 0,07x_1^2 + 0,04x_2^2 + 0,41x_1 + 1,38x_2 - 0,12x_1x_2 \quad (12)$$

As equações encontradas em (11) e (12) tem validade dentro do seguinte domínio:

Envergadura x_1	$13,20 \leq x_1 \leq 29,00$
Alongamento x_2	$15,30 \leq x_1 \leq 39,10$

Tabela 7 - Domínio de validade dos modelos

A Tabela 7 demonstra o domínio dentro do qual é possível realizar inferências.

Para o sucesso da análise da regressão, existem algumas premissas que precisam ser respeitadas. Dentre elas, destacam-se:

- As variáveis dependentes são sujeitas a erros. A média destes erros deve ser zero;
- Os erros não são correlacionados, ou seja, a matriz de variância e co-variância é diagonal e cada elemento diferente de 0 é a variância do erro.
- As variáveis dependentes precisam ser linearmente independentes;

y_1	y_2
-7,77E-11	-7,64E-12

Tabela 8 - Médias dos resíduos

Os valores da Tabela 8 indicam claramente que a primeira condição listada (a) foi satisfeita.

Quando ao item (b), analisando a Tabela 2 onde é apresentada a matriz de variância e co-variância, podemos observar que ela é diagonal e tem as variâncias dos erros representadas pelos seus termos.

O item (c) das premissas representa o principal ponto em observação neste estudo de caso.

As variáveis x_1 e x_2 , respectivamente, envergadura e alongamento, apresentam dependência linear.

Da equação (5), concluímos que existe uma correlação altamente linear entre as duas variáveis. Esta relação é conhecida por colinearidade.

A colinearidade acontece sempre que:

$$X_1 = \lambda \cdot X_2 \quad (13)$$

A equação (5) se encaixa perfeitamente neste modelo.

Na análise da regressão, quando duas ou mais variáveis explicativas são colineares, existe a multicolinearidade. A multicolinearidade pode ser dita como perfeita, quando a correlação entre duas variáveis independentes é igual a 1 ou -1.

Os indícios mais claros da multicolinearidade são:

- R_2 alto com nenhum coeficiente de regressão estatisticamente significativo segundo o teste F;
- Erros-padrão elevados;

Para o nosso caso, os valores de R^2 podem ser observados na Tabela 6 e a significância dos coeficientes de regressão podem ser consultados na Tabela 3 e Tabela 4. Percebe-se que o critério de indício (a) não é satisfeito.

Para o critério (b), percebemos erros-padrão elevados principalmente para o y_1 , chegando em casos extremos a erros maiores de 30%.

No entanto, não conseguimos afirmar de forma conclusiva que este efeito tenha sido gerado como efeito da multicolinearidade.

Por fim, concluímos que os efeitos da multicolinearidade possuem um grau de participação pequeno nesta análise, e associamos isso ao fato de que tais efeitos são diminuídos devido à quantidade amostral. Além disso, convém mencionar que multicolinearidade prejudica a interpretação das variáveis explicativas, mas não afeta a confiabilidade dos modelos obtidos.

Conclusão

A análise dos modelos encontrados demonstra que eles são representativos e podem ser utilizados como base de informação primária para suportar as decisões do projeto preliminar de um planador.

Quanto à violação de uma premissa básica, seus efeitos não se fizeram presentes devido à quantidade de amostras disponíveis para esta análise.

Referências

- BUTKEWITSH, S., 2001, Empirical Modelling of the Dynamical Behavior of Heavy Truck Side Guards Through Planned Computer Experiments.
- BUTKEWITSH, S., Notas de Aula da disciplina Estatística Aplicada ministrada no curso de Engenharia Aeronáutica e Espaço, 2008.
- JOHN D. ANDERSON, JR. Introduction to Flight. McGraw Hill, 5ª edição 2005.
- MONTGOMERY, DOUGLAS C. Introduction to Statistical Quality Control, Wiley, 6ª edição 2008.
- STELIO FRATI. The Glider, tradução por Alfred P. Scott, 2007.