

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESPECIALIZAÇÃO EM ESTATÍSTICA



PLANEJAMENTO DE EXPERIMENTO
LISTA 1

Professor: Roberto Quinino
Alunos: Meny Aparecida & Augusto Filho
Belo Horizonte - MG

Exercício 1. No livro *Planejamento e Análise de Experimentos* (Design and Analysis of Experiments), 4a. edição (John Wiley & Sons, 1988), d. C. Montgomery descreve um experimento em que um fabricante está interessado na resistência à tensão de uma fibra sintética. Suspeita-se que a resistência esteja relacionada à percentagem de algodão na fibra. Cinco níveis de percentagem de algodão são usados e cinco replicatas são corridas em uma ordem aleatória, resultando nos dados a seguir:

| Percentagem de Algodão | Observações | | | | |
|---------------------------|-------------|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 |
| 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 |

- (a) A percentagem de algodão afeta a resistência à ruptura do fio? Desenhe diagramas de caixas comparativos e faça uma análise de variância. Use $\alpha = 0,05$.

Resolução:

Utilizaremos a modelagem por variáveis indicadoras, segue a seguir, uma rápida descrição do banco de dados:

Onde atribuímos:

| % - Dummy | D1 | D2 | D3 | D4 |
|-----------|----|----|----|----|
| 15 | 1 | 0 | 0 | 0 |
| 20 | 0 | 1 | 0 | 0 |
| 25 | 0 | 0 | 1 | 0 |
| 30 | 0 | 0 | 0 | 1 |
| 35 | 1 | 1 | 1 | 1 |

Desta forma, temos:

Variáveis Indicadoras

| Perc | Obs | D1 | D2 | D3 | D4 |
|------|-----|----|----|----|----|
| 15 | 7 | 1 | 0 | 0 | 0 |
| 15 | 7 | 1 | 0 | 0 | 0 |
| . | . | . | . | . | . |
| . | . | . | . | . | . |
| . | . | . | . | . | . |
| 35 | 15 | 1 | 1 | 1 | 1 |
| 35 | 11 | 1 | 1 | 1 | 1 |

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

(i) Formulação da hipótese de interesse:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

(ii) Fixar o nível de significância do teste e o valor F tabelado:

Usaremos $\alpha = 0,05$. Trabalharemos com uma F com 4 grau de liberdade no numerador e 20 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 4 DF in numerator and 20 DF in denominator

$$P(X \leq x) \quad x \\ 0,95 \quad 2,86608$$

Logo, $F_{0,05;4,20} = 2,86608$.

(iii) Determinação da região crítica e da região de aceitação:

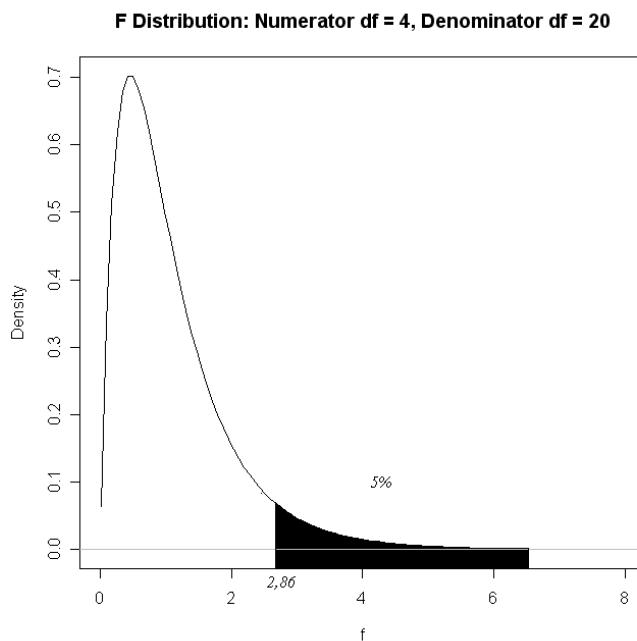


Figura 1: Determinação da região crítica.

(iv) Cálculo do valor da Estatística de Teste é dado por: $F_{cal} = \frac{\frac{R_c^2 - R_l^2}{nr}}{\frac{(1 - R_c^2)}{n - np}}$.

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 25$; $R_c^2 = 0,747$; $nr = 4$ e $np = 5$, temos:

$$F_{cal} = \frac{\frac{0,747}{4}}{\frac{(1-0,747)}{25-5}} = 14,76$$

- (v) *Conclusão: Portanto, ao nível de 5% de significância, existem evidências amostrais que nos levam a rejeição da hipótese nula, ou seja, a porcentagem de algodão afeta a resistência à ruptura do fio.*

A mesma conclusão poderia ter sido obtida pela tabela abaixo, onde os resultados a seguir apresentam a partição da soma de quadrados total e dos graus de liberdade, sendo então definida como Tabela de Análise de Variância:

Analysis of Variance

| Source | DF | SS | MS | F | P |
|----------------|----|--------|--------|-------|-------|
| Regression | 4 | 475,76 | 118,94 | 14,76 | 0,000 |
| Residual Error | 20 | 161,20 | 8,06 | | |
| Total | 24 | 636,96 | | | |

Logo, pela análise de variância acima, rejeitamos da mesma forma a hipótese nula ao nível de 5% de significância.

A seguir, apresentaremos o box-plot comparativo:

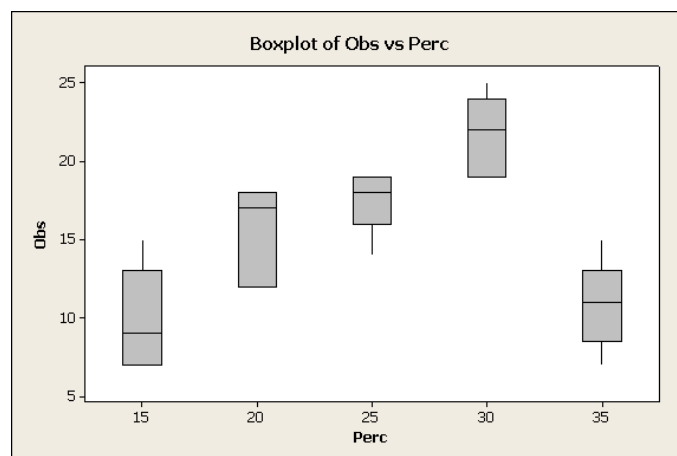


Figura 2: Diagrama de caixas comparativos

Através do box-plot acima, podemos concluir que a porcentagem de algodão afeta a resistência à ruptura do algodão.

(b) Plote a resistência média à tensão contra a % de algodão e interprete os resultados;

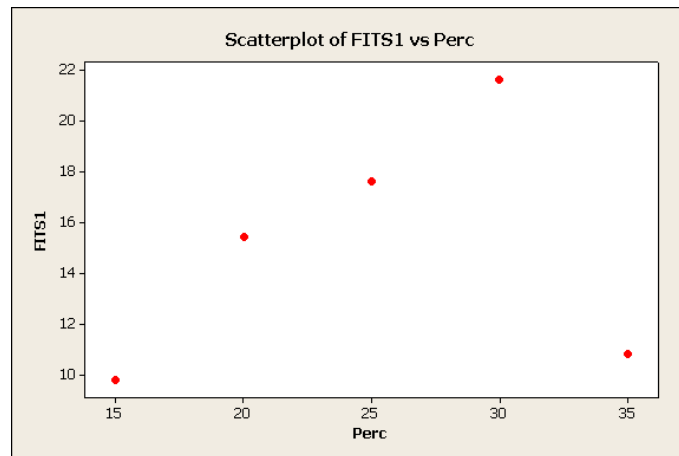


Figura 3: Diagrama de caixas comparativos

Através do diagrama de dispersão acima é possível perceber que a resistência média a tensão (Fits) contra a percentagem de algodão aumenta em 30% e decai em 35%.

(c) Encontre os resíduos e examine-os em relação à inadequação do modelo.

Resolução:

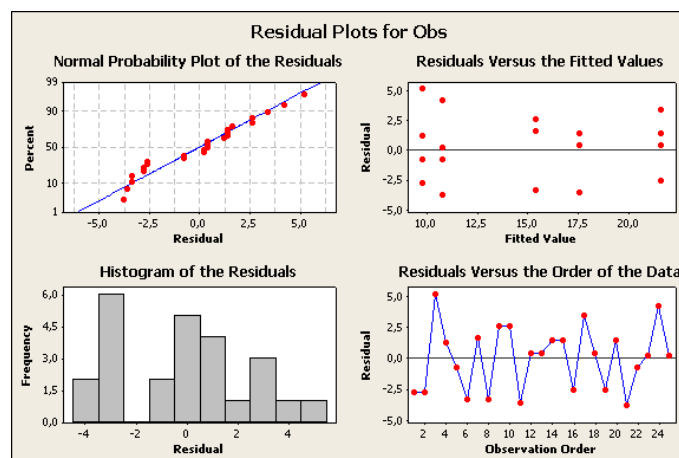


Figura 4: Gráficos de resíduos

Em relação aos gráfico de probabilidade normal observou-se que os pontos centrais estão localizados, de forma aproximada, ao longo de uma reta, o que indica que os componentes de erro do modelo seguem uma distribuição normal, o gráfico da ordem versus resíduo não mostra nenhuma configuração especial que revelasse uma associação dos resíduos com a ordem de coleta, o que indica independência e o gráfico dos Fits versus resíduo mostra que as faixas de dispersão dos resíduos possuem comprimentos aproximadamente iguais, o que indica que a suposição de igualdade de variância é válida.

Exercício 2. No artigo “Planejamento Ortogonal para Otimização de Processo e Sua Aplicação a Ataque por Plasma” (Orthogonal Design for Process Optimization and Its Application to Plasma Etching), no periódico *Solid State Technology*, maio de 1987, G. Z. Yin e d. W. Jilie descrevem um experimento para determinar o efeito da taxa de escoamento de C_2F_6 sobre a uniformidade do ataque químico em uma pastilha de silicone usada na fabricação de um circuito integrado. Três taxas de escoamento são usadas no experimento e a unidade (%) resultante, para seis replicatas é mostrada a seguir:

| Escoamento de C_2F_6 | Observações | | | | | |
|---------------------------|-------------|-----|-----|-----|-----|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| 125 | 2,7 | 4,6 | 2,6 | 3,0 | 3,2 | 3,8 |
| 160 | 4,9 | 4,6 | 5,0 | 4,2 | 3,6 | 4,2 |
| 200 | 4,6 | 3,4 | 2,9 | 3,5 | 4,1 | 5,1 |

(a) A taxa de escoamento de C_2F_6 afeta a uniformidade do ataque químico?

Resolução:

Utilizaremos a modelagem por variáveis dummy, a seguir, uma rápida descrição do banco de dados:

Onde atribuímos:

| <i>Esc - Dummy</i> | <i>D1</i> | <i>D2</i> |
|--------------------|-----------|-----------|
| 125 | 1 | 0 |
| 160 | 0 | 1 |
| 200 | 1 | 1 |

Desta forma, temos:

Variáveis Indicadoras

| <i>Escoam</i> | <i>Obs</i> | <i>D1</i> | <i>D2</i> |
|---------------|------------|-----------|-----------|
| 125 | 2,7 | 1 | 0 |
| 125 | 4,6 | 1 | 0 |
| 125 | 2,6 | 1 | 0 |
| . | . | . | . |
| . | . | . | . |
| . | . | . | . |
| 200 | 3,5 | 1 | 1 |
| 200 | 4,1 | 1 | 1 |
| 200 | 5,1 | 1 | 1 |

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

(i) Formulação da hipótese de interesse:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

(ii) Fixar o nível de significância do teste e o valor F tabelado:

Usaremos $\alpha = 0,05$. Trabalharemos com uma F com 2 grau de liberdade no numerador e 15 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 2 DF in numerator and 15 DF in denominator

$$P(X \leq x) \quad x \\ 0,95 \quad 3,68232$$

Logo, $F_{0,05;2,15} = 3,68232$.

(iii) Determinação da região crítica e da região de aceitação:

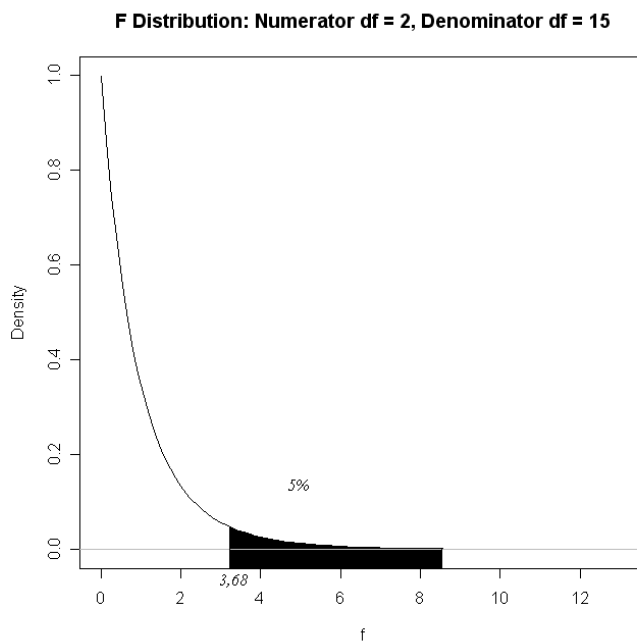


Figura 5: Determinação da região crítica.

(iv) Cálculo do valor da Estatística de Teste é dado por: $F_{cal} = \frac{\frac{R_c^2 - R_l^2}{nr}}{\frac{(1 - R_c^2)}{n - np}}$.

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 18$; $R_c^2 = 0,323$; $nr = 2$ e $np = 3$, temos:

$$F_{cal} = \frac{\frac{0,323}{2}}{\frac{(1-0,323)}{18-3}} = 3,59$$

- (v) *Conclusão: Portanto, ao nível de 5% de significância, existem evidências amostrais que nos levam a **não** rejeição da hipótese nula, ou seja, a taxa de escoamento de C_2F_6 **não afeta** a uniformidade do ataque químico em uma pastilha de silicone usado na fabricação de um circuito integrado.*

A mesma conclusão poderia ter sido obtida pela tabela abaixo, onde os resultados a seguir apresentam a partição da soma de quadrados total e dos graus de liberdade, sendo então definida como Tabela de Análise de Variância:

Analysis of Variance

| Source | DF | SS | MS | F | P |
|----------------|----|---------|--------|------|-------|
| Regression | 2 | 3,6478 | 1,8239 | 3,59 | 0,053 |
| Residual Error | 15 | 7,6300 | 0,5087 | | |
| Total | 17 | 11,2778 | | | |

Logo, pela análise de variância acima, **não rejeitamos** da mesma forma a hipótese nula ao nível de 5% de significância.

- (b) *Os resíduos indicam qualquer problema com as suposições propostas?* **Resolução:**

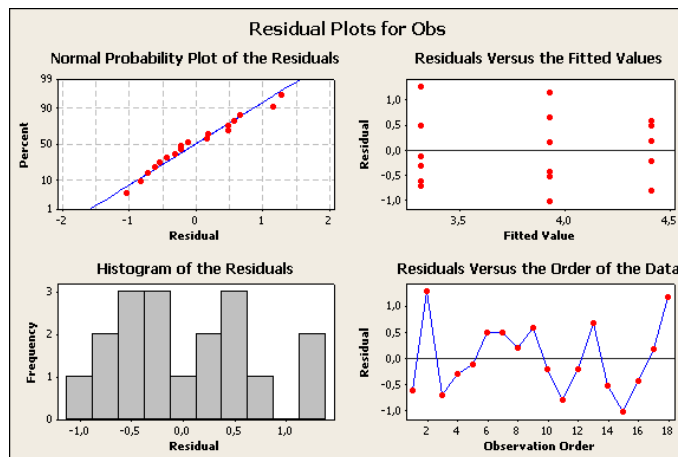


Figura 6: Gráfico dos resíduos.

Os gráficos acima não mostram indícios de falta de adequação do modelo e das suposições a ele associadas, logo as conclusões obtidas por meio destes testes são válidas.

Exercício 3. A resistência do concreto à compressão está sendo estudada e quatro técnicas diferentes de mistura estão sendo investigadas. Os seguintes dados foram coletados.

| <i>Técnica de Mistura</i> | <i>Resistência à Compressão (psi)</i> | | | |
|---------------------------|---------------------------------------|------|------|------|
| 1 | 3129 | 3000 | 2865 | 2890 |
| 2 | 3200 | 3300 | 2975 | 3150 |
| 3 | 2800 | 2900 | 2985 | 3050 |
| 4 | 2600 | 2700 | 2600 | 2765 |

- (a) Teste a hipótese de que as técnicas de mistura afetam a resistência do concreto. Use $\alpha = 0,05$.

Resolução:

Utilizaremos a modelagem por variáveis dummy, a seguir, uma rápida descrição do banco de dados:

Onde atribuímos:

| <i>Mist - Dummy</i> | <i>D1</i> | <i>D2</i> | <i>D3</i> |
|---------------------|-----------|-----------|-----------|
| 1 | 1 | 0 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 1 | 1 | 1 |

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

- (i) Formulação da hipótese de interesse:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \beta_3 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

- (ii) Fixar o nível de significância do teste e o valor F tabelado:

Usaremos $\alpha = 0,05$. Trabalharemos com uma F com 3 grau de liberdade no numerador e 12 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 3 DF in numerator and 12 DF in denominator

$P(X \leq x) \quad x$

0,95 3,49029

Logo, $F_{0,05;3,12} = 3,49029$.

(iii) *Determinação da região crítica e da região de aceitação:*

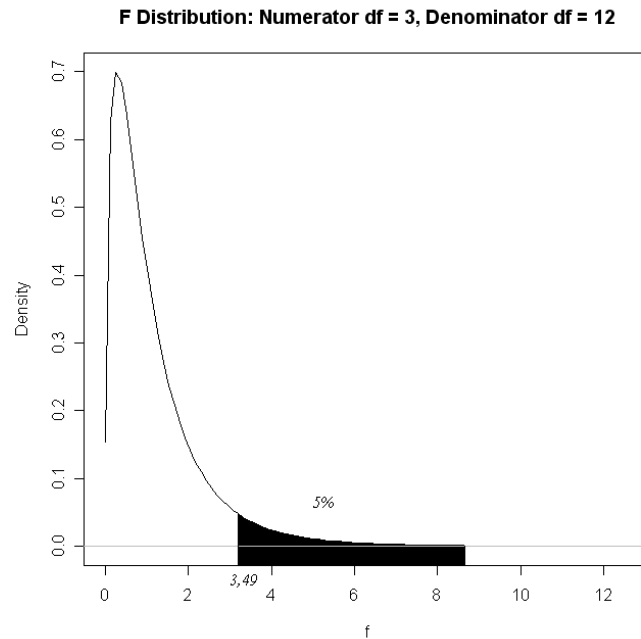


Figura 7: Determinação da região crítica.

(iv) *Cálculo do valor da Estatística de Teste é dado por:* $F_{cal} = \frac{R_c^2 - R_I^2}{\frac{nr}{n-np} (1-R_c^2)}$.

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 16$; $R_c^2 = 0,761$; $nr = 3$ e $np = 4$, temos:

$$F_{cal} = \frac{0,761}{\frac{3}{(1-0,761)} \cdot \frac{3}{16-4}} = 12,74$$

(v) *Conclusão: Portanto, ao nível de 5% de significância, existem evidências amostrais que nos levam a rejeição da hipótese nula, ou seja, as técnicas de mistura afetam a resistência do concreto a um nível de 5% de significância.*

A mesma conclusão poderia ter sido obtida pela ANOVA abaixo, onde os resultados a seguir apresentam a partição da soma de quadrados total e dos graus de liberdade:

Analysis of Variance

| Source | DF | SS | MS | F | P |
|----------------|----|--------|--------|-------|-------|
| Regression | 3 | 489740 | 163247 | 12,73 | 0,000 |
| Residual Error | 12 | 153908 | 12826 | | |
| Total | 15 | 643648 | | | |

(b) Encontre o valor P para a estatística F calculada no item (a).

Resolução:

Através do Minitab, o p -valor é dado como o complementar do resultado seguinte:

Cumulative Distribution Function

F distribution with 3 DF in numerator and 12 DF in denominator

| | |
|-------|-----------------|
| x | $P(X \leq x)$ |
| 12,73 | 0,999512 |

Logo o p -valor é igual a: $1 - P(X \leq 12,73) = 1 - 0,999512 = 0,000488$.

Como visto na ANOVA acima.

(c) Analise os resíduos provenientes desse experimento.

Resolução:

A seguir, temos o gráfico de resíduos para a variável em estudo:

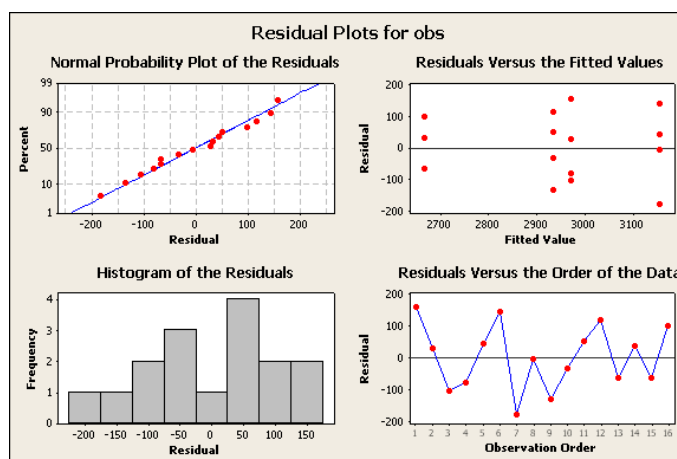


Figura 8: Gráfico de resíduos.

Observando o gráfico de resíduos contra a ordem de coleta dos dados não é possível notar quaisquer configurações especiais que revelassem uma associação dos resíduos com a ordem de coleta das observações, o que indica que a suposição de independência não parece ter sido violada. O gráfico de resíduos contra médias (Fitted) mostra que as faixas de dispersão dos resíduos em cada média (tratamento) tem comprimentos aproximadamente iguais, o que indica que a suposição de igualdade de variância pode ser considerada válida. Já o gráfico de probabilidade normal nos mostra que os pontos estão distribuídos ao longo de uma reta, o que indica que os componentes de erro do modelo seguem uma distribuição normal, portanto, como não foram encontrados indícios de falta de adequação do modelo e das suposições a ele associada, as conclusões obtidas por meio dos testes podem ser consideradas válidas.

Exercício 4. *Um experimento foi feito para determinar se quatro temperaturas específicas de queima afetam a densidade de um certo tipo de tijolo. O experimento conduziu aos seguintes dados:*

| Temperatura (F) | Densidade | | | | | | |
|--------------------|-----------|------|------|------|------|------|------|
| | 100 | 21,8 | 21,9 | 21,7 | 21,6 | 21,7 | 21,5 |
| 125 | 21,7 | 21,4 | 21,5 | 21,5 | – | – | – |
| 150 | 21,9 | 21,8 | 21,8 | 21,6 | 21,5 | – | – |
| 175 | 21,9 | 21,7 | 21,8 | 21,7 | 21,6 | 21,8 | – |

(a) *A temperatura de queima afeta a densidade dos tijolos? Use $\alpha = 0,05$;*

Resolução:

Utilizaremos a modelagem por variáveis dummy, a seguir, uma rápida descrição do banco de dados:

Onde atribuímos:

| Temp - Dummy | D1 | D2 | D3 |
|--------------|----|----|----|
| 100 | 1 | 0 | 0 |
| 125 | 0 | 1 | 0 |
| 150 | 0 | 0 | 1 |
| 175 | 1 | 1 | 1 |

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

(i) *Formulação da hipótese de interesse:*

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \beta_3 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

(ii) *Fixar o nível de significância do teste e o valor F tabelado:*

Usaremos $\alpha = 0,05$. Trabalharemos com uma F com 3 grau de liberdade no numerador e 18 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 3 DF in numerator and 18 DF in denominator

$$P(X \leq x) \quad x \\ 0,95 \quad 3,15991$$

Logo, $F_{0,05;3,18} = 3,15991$.

(iii) *Determinação da região crítica e da região de aceitação:*

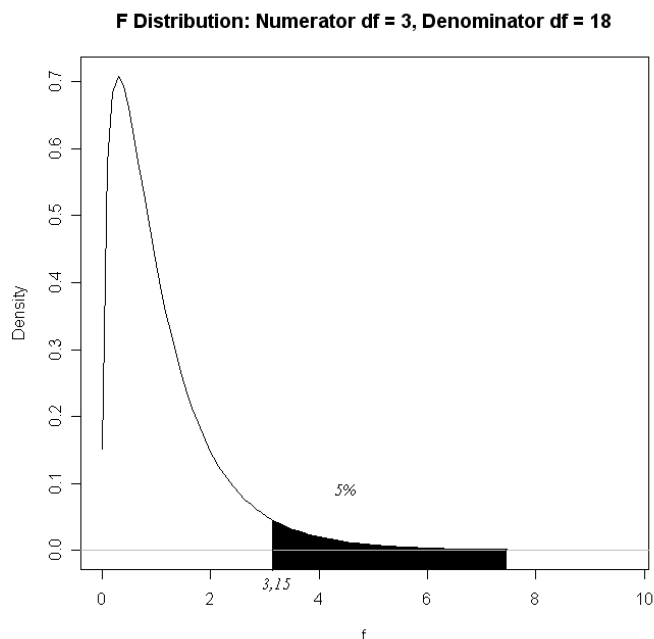


Figura 9: Determinação da região crítica.

(iv) *Cálculo do valor da Estatística de Teste é dado por:* $F_{cal} = \frac{R_c^2 - R_I^2}{\frac{nr}{n-np} \cdot (1-R_c^2)}$.

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 22$; $R_c^2 = 0,304$; $nr = 3$ e $np = 4$, temos:

$$F_{cal} = \frac{0,304}{\frac{3}{(1-0,304)} \cdot \frac{22-4}{22-4}} = 2,62$$

(v) *Conclusão: Portanto, ao nível de 5% de significância, existem evidências amostrais que nos levam a **não** rejeição da hipótese nula, ou seja, a temperatura de queima não afeta a densidade dos tijolos a um nível de significância de 5%.*

A mesma conclusão poderia ter sido obtida pela ANOVA abaixo, onde os resultados a seguir apresentam a partição da soma de quadrados total e dos graus de liberdade:

Analysis of Variance

| Source | DF | SS | MS | F | P |
|----------------|----|---------|---------|------|-------|
| Regression | 3 | 0,13911 | 0,04637 | 2,62 | 0,083 |
| Residual Error | 18 | 0,31907 | 0,01773 | | |
| Total | 21 | 0,45818 | | | |

- (b) Encontre o valor p para a estatística F calculada no item (a);

Resolução:

Através do Minitab, o p -valor é dado como o complementar do resultado seguinte:

Cumulative Distribution Function

F distribution with 3 DF in numerator and 18 DF in denominator

$$x \quad P(X \leq x)$$

$$2,62 \quad 0,917664$$

Logo o p -valor é igual a: $1 - P(X \leq 2,62) = 1 - 0,917664 = 0,082336$.

Como visto na ANOVA acima.

- (c) Analise os resíduos provenientes do experimento.

Resolução:

A seguir, temos o gráfico de resíduos para a variável em estudo:

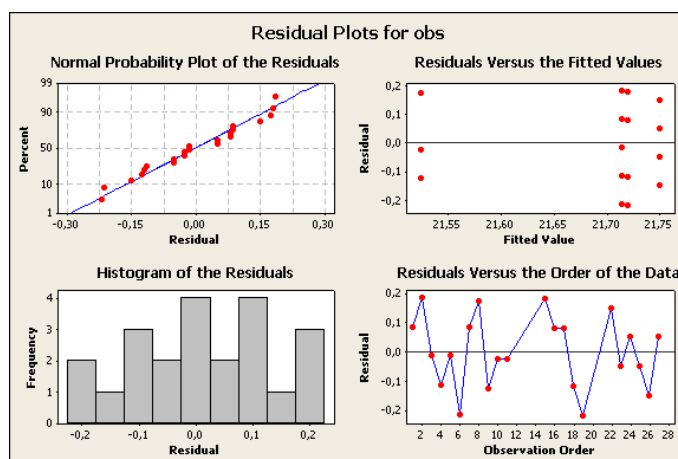


Figura 10: Gráfico de resíduos.

O gráfico de resíduos contra a ordem de coleta dos dados não mostrou qualquer configuração especial que revelasse uma associação dos resíduos com a ordem de coleta das observações, o que indica que a suposição de independência não parece ter sido violada. O gráfico de resíduos contra médias (Fitted) mostra as faixas de dispersão dos resíduos com comprimentos aproximadamente iguais, o que indica que a suposição de igualdade de variância pode ser considerada válida. E o gráfico de probabilidade normal mostra os pontos distribuídos ao longo da reta, o que indica que os componentes de erro do modelo seguem uma distribuição normal. Portanto como não foram encontrados indícios de falta de ajustamento, as conclusões obtidas nos testes podem ser consideradas válidas.

Exercício 5. Um engenheiro eletrônico está interessado no efeito, na condutividade do tubo, de cinco tipos diferentes de revestimento de tubos de raios catódicos em uma tela de um sistema de telecomunicações. Os seguintes dados de condutividade são obtidos.

| Tipo de Revestimento | Condutividade | | | |
|-------------------------|---------------|-----|-----|-----|
| | 1 | 143 | 141 | 150 |
| 2 | 152 | 149 | 137 | 143 |
| 3 | 134 | 133 | 132 | 127 |
| 4 | 129 | 127 | 132 | 129 |
| 5 | 147 | 148 | 144 | 142 |

(a) Há qualquer diferença na condutividade devido ao tipo de revestimento? Use $\alpha = 0,01$.

Resolução:

Utilizaremos a modelagem por variáveis dummy, a seguir, uma rápida descrição da codificação utilizada no banco de dados:

Onde atribuímos:

| Recob - Dummy | D1 | D2 | D3 | D4 |
|---------------|----|----|----|----|
| 1 | 1 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 0 |
| 3 | 0 | 0 | 1 | 0 |
| 4 | 0 | 0 | 0 | 1 |
| 5 | 1 | 1 | 1 | 1 |

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

(i) Formulação da hipótese de interesse:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

(ii) Fixar o nível de significância do teste e o valor F tabelado:

Usaremos $\alpha = 0,01$. Trabalharemos com uma F com 4 grau de liberdade no numerador e 15 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 4 DF in numerator and 15 DF in denominator

$$P(X \leq x) = 0,99 \quad x = 4,89321$$

Logo, $F_{0,01;4,15} = 4,89321$.

(iii) Determinação da região crítica e da região de aceitação:

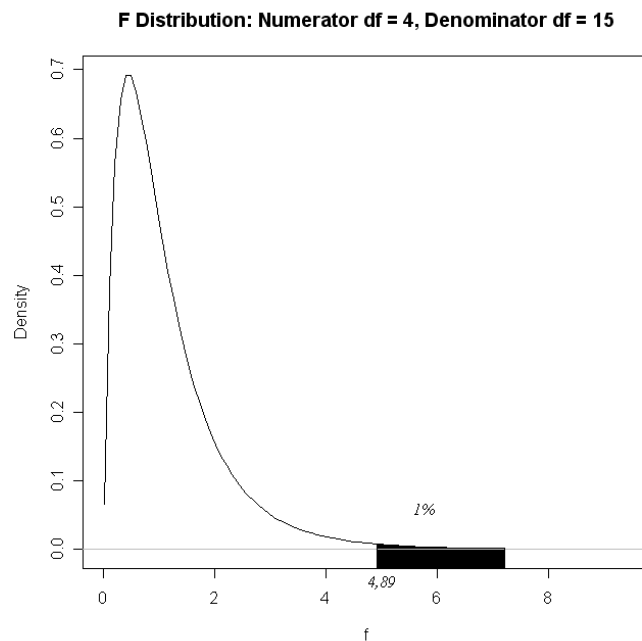


Figura 11: Determinação da região crítica.

(iv) Cálculo do valor da Estatística de Teste é dado por: $F_{cal} = \frac{R_c^2 - R_I^2}{\frac{nr}{n-np}}$.

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 20$; $R_c^2 = 0,813$; $nr = 4$ e $np = 5$, temos:

$$F_{cal} = \frac{0,813}{\frac{4}{20-5}} = 16,30$$

(v) Conclusão: Portanto, ao nível de 1% de significância, existem evidências amostrais que nos levam a rejeição da hipótese nula, ou seja, existe diferença na condutividade devido ao tipo de recobrimento a um nível de 1% de significância.

(b) Analise os resíduos provenientes desse experimento;

Resolução:

A seguir, temos o gráfico de resíduos para a variável em estudo:

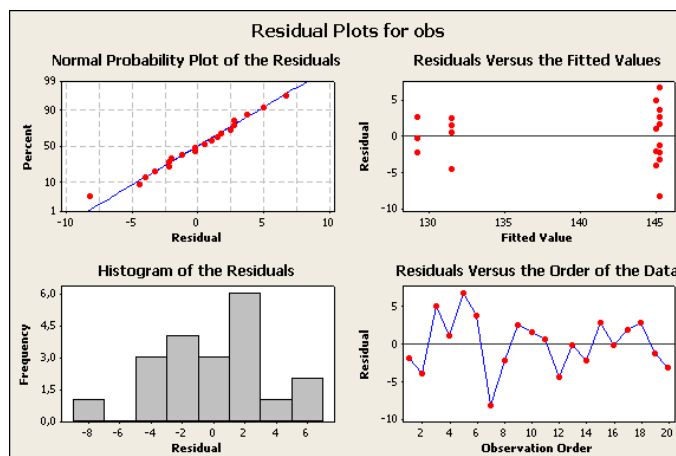


Figura 12: Gráfico de resíduos.

Observando o gráfico de resíduos contra a ordem de coleta dos dados não foi possível observar quaisquer configurações especiais que revelassem uma associação dos resíduos com a ordem de coleta das observações, o que indicou que a suposição de independência não parecia ter sido violada. O gráfico de resíduos contra médias nos mostra que as faixas de dispersão dos resíduos em cada média (tratamento) possuem comprimentos aproximadamente iguais, o que indica que a suposição de igualdade de variâncias poderia ser considerada válida e o gráfico de probabilidade normal nos mostra que os pontos estão distribuídos ao longo de uma reta, o que nos indica que os componentes de erro do modelo segue uma distribuição normal. Portanto, como não foram encontrados indícios de falta de adequação do modelo e das suposições a ele associadas, as conclusões obtidas por meio dos testes acima podem ser consideradas válidas.

(c) Construa uma estimativa do intervalo de 95% para a média do tipo 1 de recobrimento. Construa uma estimativa do intervalo de 99% para a diferença média entre os tipos 1 e 4 de recobrimento.

Resolução:

Abaixo, temos a construção da estimativa do intervalo de 95% para a média do tipo 1 de recobrimento.

$$\bar{y}_i \cdot -t_{\alpha/2, a(n-1)} \cdot \sqrt{\frac{MQE}{n}} \leq \mu_i \leq \bar{y}_i \cdot +t_{\alpha/2, a(n-1)} \cdot \sqrt{\frac{MQE}{n}}$$

Temos: $\bar{y}_1 = 145$ $t_{0,025;15} = 2,1315$ $MQ_E = 16,22$

$$\left[\bar{y}_1 \pm t_{0,025;15} \cdot \sqrt{\frac{MQ_E}{n}} \right] = \left[145 \pm 2,1315 \cdot \sqrt{\frac{16,22}{4}} \right] = [140,7; 149,3]$$

Desse modo, o intervalo de confiança desejado é:

$$140,7 \leq \mu_1 \leq 149,3$$

A seguir, temos a construção da estimativa intervalar para a diferença entre média entre os tipos 1 e 4 de revestimento.

Utilizaremos o seguinte procedimento para encontrar o intervalo para a diferença entre as médias dos dois tipos de revestimento.

$$\mu_i - \mu_j = (\bar{y}_i - \bar{y}_j) \pm t_{\alpha/2; k(n-1)} \sqrt{\frac{2 \cdot QM_{Res}}{n}}$$

Onde $t_{\alpha; k(n-1)}$ é o valor tabelado da distribuição *t*-Student com $k(n-1)$ graus de liberdade e $QM_{Res} = QM_{erro}$

Resolução:

$$\left[15,75 \pm t_{0,005;15} \cdot \sqrt{\frac{2 \cdot 16,22}{4}} \right] = [15,75 \pm 2,9467 \cdot 2,85] = [7,36; 24,15]$$

Logo, o intervalo de confiança de 99% para $\mu_1 - \mu_4$ é:

$$7,36 \leq \mu_1 - \mu_4 \leq 24,15$$

Uma vez que o intervalo de confiança não inclui o zero, concluímos que há evidências de diferenças na condutividade devido ao tipo de revestimento ao nível de 1% de significância.

Exercício 6. O tempo de resposta em milissegundos foi determinado para três tipos diferentes de circuitos em uma calculadora eletrônica. Os resultados estão registrados a seguir:

| Tipo de Circuito | Resposta | | | | |
|------------------|----------|----|----|----|----|
| | 1 | 19 | 22 | 20 | 18 |
| 2 | 20 | 21 | 33 | 27 | 40 |
| 3 | 16 | 15 | 18 | 26 | 17 |

- (a) Usando $\alpha = 0,01$, teste a hipótese de os três tipos de circuito terem o mesmo tempo de resposta;

Resolução:

Utilizaremos a modelagem por variáveis dummy, a seguir, uma rápida descrição da codificação utilizada no banco de dados:

Onde atribuímos:

| Tipo - Dummy | D1 | D2 |
|--------------|----|----|
| 1 | 1 | 0 |
| 2 | 0 | 1 |
| 3 | 1 | 1 |

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

- (i) Formulação da hipótese de interesse:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

- (ii) Fixar o nível de significância do teste e o valor F tabelado:

Usaremos $\alpha = 0,01$. Trabalharemos com uma F com 2 grau de liberdade no numerador e 12 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 2 DF in numerator and 12 DF in denominator

$$P(X \leq x) \quad x$$

$$0,99 \quad 6,92661$$

Logo, $F_{0,01;2,12} = 6,92661$.

(iii) *Determinação da região crítica e da região de aceitação:*

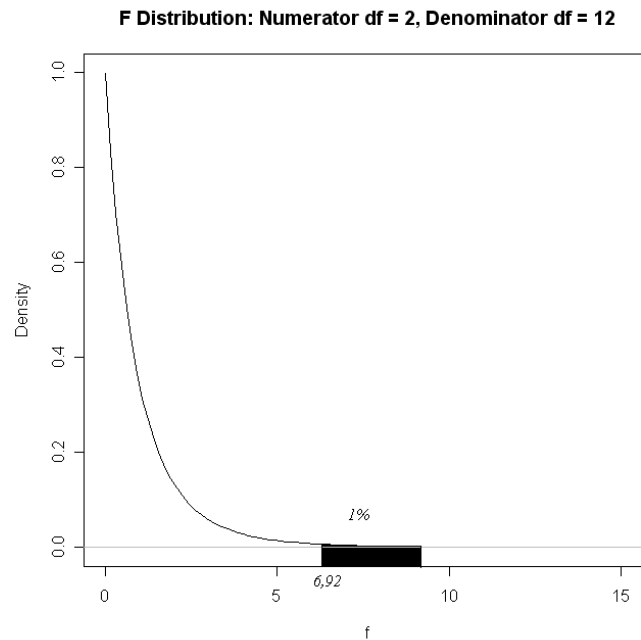


Figura 13: Determinação da região crítica.

(iv) *Cálculo do valor da Estatística de Teste é dado por:* $F_{cal} = \frac{\frac{R_c^2 - R_I^2}{nr}}{\frac{(1 - R_c^2)}{n - np}}$.

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 15$; $R_c^2 = 0,40$; $nr = 2$ e $np = 3$, temos:

$$F_{cal} = \frac{\frac{0,40}{2}}{\frac{(1-0,40)}{15-3}} = 4,01$$

(v) *Conclusão: Portanto, ao nível de 1% de significância, existem evidências amostrais que nos levam a **não** rejeição da hipótese nula, ou seja, não existe diferença entre os três tempos de respostas dos circuitos ao nível de 1% de significância.*

(b) Analise os resíduos provenientes desse experimento;

Resolução:

A seguir, temos o gráfico de resíduos para a variável em estudo:

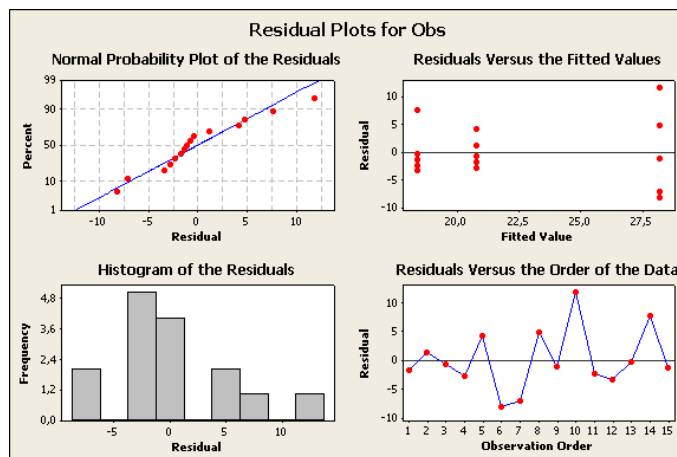


Figura 14: Gráfico de resíduos.

Através de uma análise atenta nos gráficos de resíduos, não foram encontrados indícios de falta de adequação do modelo e das suposições a ele associadas, as conclusões obtidas por meio dos testes realizados podem ser consideradas válidas.

(c) Encontre uma estimativa do intervalo de confiança de 95% para o tempo de resposta, no caso do circuito de número três.

Resposta:

Utilizaremos a seguinte fórmula:

$$\bar{y}_i \cdot -t_{\alpha/2, a(n-1)} \cdot \sqrt{\frac{MQ_E}{n}} \leq \mu_i \leq \bar{y}_i \cdot +t_{\alpha/2, a(n-1)} \cdot \sqrt{\frac{MQ_E}{n}}$$

$$\text{Onde: } \bar{y}_{1\cdot} = 20,8 \quad t_{0,025;12} = 2,1788 \quad MQ_E = 32,57$$

$$\left[\bar{y}_{1\cdot} \pm t_{0,025;12} \cdot \sqrt{\frac{MQ_E}{n}} \right] = \left[20,8 \pm 2,1788 \cdot \sqrt{\frac{32,57}{5}} \right] = [15,24; 26,36]$$

Desse modo, o intervalo de confiança desejado é:

$$15,24 \leq \mu_1 \leq 26,36$$

Logo, o intervalo de confiança para a verdadeira média populacional do tipo de circuito do Tipo 3 é de 15,24 a 26,36 ao nível de 95% de confiança.

Exercício 7. Um artigo na revista *ACI Materials Journal* (Vol. 84, 1987, pp. 213-216) descreve vários experimentos investigando a vibração da massa de concreto a fim de remover o ar preso. Um cilindro de 3" x 6" foi usado e o número de vezes em que esse bastão foi usado é a variável de planejamento do corpo de prova de concreto. Os dados são mostrados na seguinte tabela:

| Nível de Giro | Resistência à Compressão | | |
|---------------|--------------------------|------|------|
| | 10 | 1530 | 1530 |
| 15 | 1610 | 1650 | 1500 |
| 20 | 1560 | 1730 | 1530 |
| 25 | 1500 | 1490 | 1510 |

(a) Há alguma diferença na resistência compressiva devido ao nível de vibração;

Resolução:

Utilizaremos a modelagem por variáveis dummy, a seguir, uma rápida descrição da codificação utilizada no banco de dados:

Onde atribuímos:

| Nível - Dummy | D1 | D2 | D3 |
|---------------|----|----|----|
| 10 | 1 | 0 | 0 |
| 15 | 0 | 1 | 0 |
| 20 | 0 | 0 | 1 |
| 25 | 1 | 1 | 1 |

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

(i) Formulação da hipótese de interesse:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \beta_3 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

(ii) Fixar o nível de significância do teste e o valor F tabelado:

Usaremos $\alpha = 0,05$. Trabalharemos com uma F com 3 grau de liberdade no numerador e 8 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 3 DF in numerator and 8 DF in denominator

$$P(X \leq x) = 0,95 \quad x = 4,06618$$

$$\text{Logo, } F_{0,05;3,8} = 4,06618.$$

(iii) *Determinação da região crítica e da região de aceitação:*

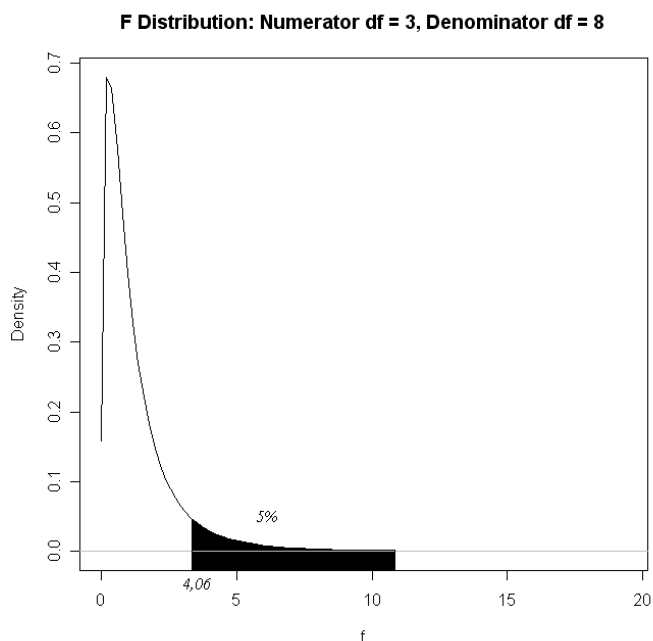


Figura 15: Determinação da região crítica.

(iv) *Cálculo do valor da Estatística de Teste é dado por:* $F_{cal} = \frac{R_c^2 - R_I^2}{\frac{nr}{(1-R_c^2)} \cdot \frac{1}{n-np}}$

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 12$; $R_c^2 = 0,412$; $nr = 3$ e $np = 4$, temos:

$$F_{cal} = \frac{\frac{0,412}{3}}{\frac{(1-0,412)}{12-4}} = 1,87$$

(v) *Conclusão: Portanto, ao nível de 5% de significância, existem evidências amostrais que nos levam a **não** rejeição da hipótese nula, ou seja, não existe diferença na resistência compressiva devido ao nível de vibração ao nível de 5% de significância.*

(b) *Encontre o valor P para a estatística F calculada no item (a);*

A seguir, veremos uma saída computacional pelo software Minitab.

Resolução:

Através do Minitab, o p -valor é dado como o complementar do resultado seguinte:

Cumulative Distribution Function

F distribution with 3 DF in numerator and 8 DF in denominator

x $P(X \leq x)$
 1,87 0,786991

Logo o p -valor é igual a: $1 - P(X \leq 1,87) = 1 - 0,786991 = 0,213009$.

- (c) Analise os resíduos provenientes do experimento. Que conclusões você pode tirar acerca das suposições propostas no modelo?

Resolução:

A seguir, temos o gráfico de resíduos para a variável em estudo:

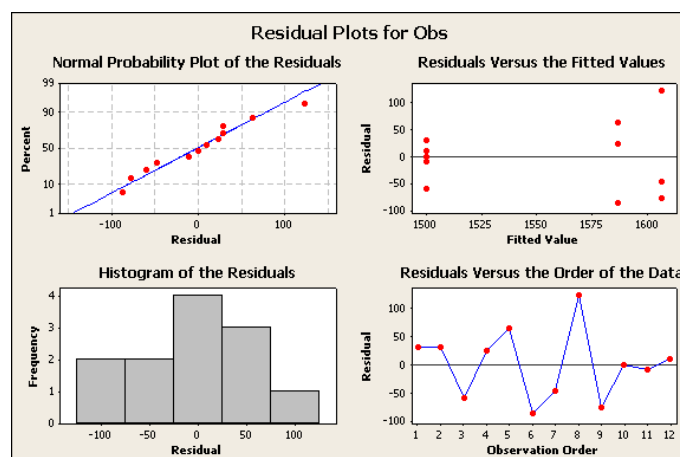


Figura 16: Gráfico de resíduos.

O gráfico dos resíduos não indicou falta de adequação do modelo e das suposições a ele associadas, as conclusões obtidas por meio dos testes realizados podem ser consideradas válidas.

Exercício 8. Um artigo na revista *Environment International* (Vol. 18, No. 4, 1992) descreve um experimento em que se investigou a quantidade de radônio liberado em chuveiros. água enriquecida com radônio foi usada no experimento e seis tipos diferentes de diâmetros do orifício foram testados nas cabeças do chuveiro. Os dados do experimento são mostrados na seguinte tabela.

| Diâmetro do Orifício | Radônio Liberado (%) | | | |
|----------------------|----------------------|----|----|----|
| | 0,37 | 80 | 83 | 83 |
| 0,51 | 75 | 75 | 79 | 79 |
| 0,71 | 74 | 73 | 76 | 77 |
| 1,02 | 67 | 72 | 74 | 74 |
| 1,40 | 62 | 62 | 67 | 69 |
| 1,99 | 60 | 61 | 64 | 66 |

(a) O tamanho do orifício afeta a percentagem média de radônio liberado? Use $\alpha = 0,05$.

Resolução:

Utilizaremos a modelagem por variáveis dummy, a seguir, uma rápida descrição da codificação utilizada no banco de dados:

Onde atribuímos:

| Diamet - Dummy | D1 | D2 | D3 | D4 | D5 |
|----------------|----|----|----|----|----|
| 0,37 | 1 | 0 | 0 | 0 | 0 |
| 0,51 | 0 | 1 | 0 | 0 | 0 |
| 0,71 | 0 | 0 | 1 | 0 | 0 |
| 1,02 | 0 | 0 | 0 | 1 | 0 |
| 1,40 | 0 | 0 | 0 | 0 | 1 |
| 1,99 | 1 | 1 | 1 | 1 | 1 |

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

(i) Formulação da hipótese de interesse:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

(ii) Fixar o nível de significância do teste e o valor F tabelado:

Usaremos $\alpha = 0,05$. Trabalharemos com uma F com 5 grau de liberdade no numerador e 18 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 5 DF in numerator and 18 DF in denominator

$$P(X \leq x) = 0,95 \quad x = 2,77285$$

Logo, temos: $F_{0,05;5,18} = 2,77285$.

(iii) Determinação da região crítica e da região de aceitação:

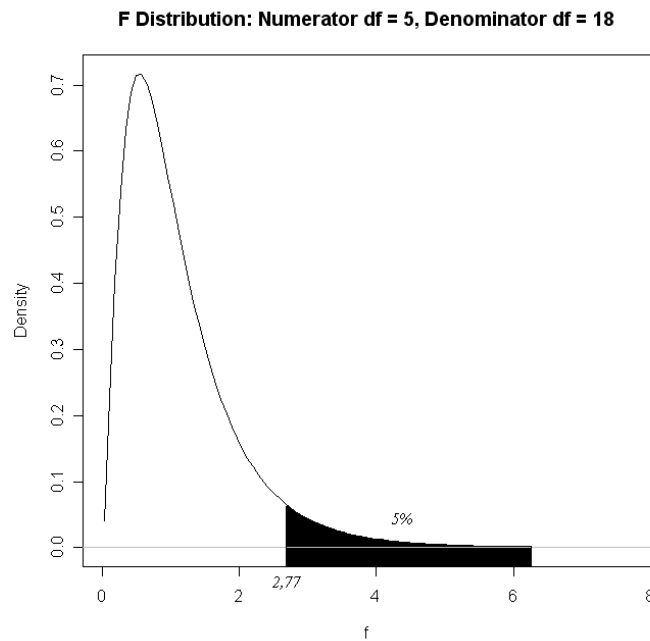


Figura 17: Determinação da região crítica.

(iv) Cálculo do valor da Estatística de Teste é dado por: $F_{cal} = \frac{R_c^2 - R_I^2}{\frac{nr}{(1-R_c^2)} \cdot \frac{1}{n-np}}$.

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 24$; $R_c^2 = 0,896$; $nr = 5$ e $np = 6$, temos:

$$F_{cal} = \frac{\frac{0,896}{5}}{\frac{(1-0,896)}{24-6}} = 30,85$$

(v) Conclusão: Portanto, ao nível de 5% de significância, existem fortes evidências amostrais que nos levam a rejeição da hipótese nula, ou seja, o tamanho do orifício afeta a percentagem média de radônio liberado ao nível de 5% de significância.

- (b) Encontre o valor P para a estatística F calculada no item (a).

Resolução:

Através do Minitab, o p -valor é dado como o complementar do resultado seguinte:

Cumulative Distribution Function

F distribution with 5 DF in numerator and 18 DF in denominator

| | |
|-------|-----------------|
| x | $P(X \leq x)$ |
| 30,85 | 1,00000 |

Logo o p -valor é igual a: $1 - P(X \leq 30,85) = 1 - 1,0000 = 0,0000$.

- (c) Analise os resíduos provenientes desse experimento;

Resolução:

A seguir, temos o gráfico de resíduos para a variável em estudo:

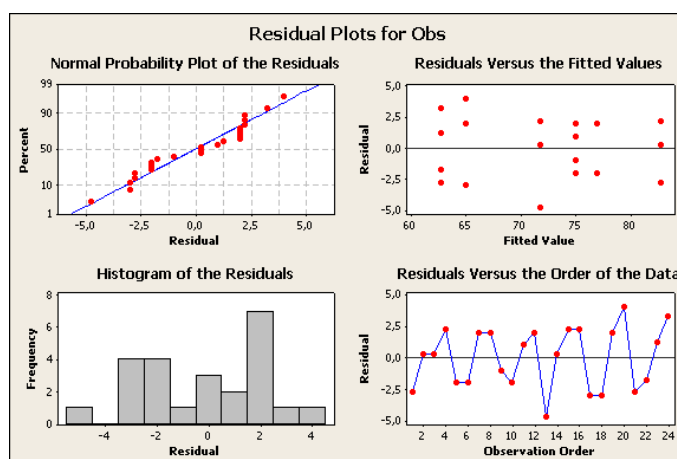


Figura 18: Gráfico de resíduos.

Não foram encontradas indícios de falta de adequação das suposições de regressão, logo as conclusões utilizando esse modelo podem ser consideradas válidas.

- (d) Encontre um intervalo de confiança de 95% para a percentagem média de radônio liberado, quando o diâmetro do orifício for 1,40;

Resolução:

Temos: $\bar{y}_1 = 65$ $t_{0,025;18} = 2,1009$ $MQ_E = 7,35$

$$\left[\bar{y}_1 \pm t_{0,025;18} \cdot \sqrt{\frac{MQ_E}{n}} \right] = [62,99; 67,01]$$

Exercício 9. Um trabalho no periódico Journal of the Association of Asphalt Paving Technologists (Vol. 59,1990) descreve um experimento com o objetivo de determinar o efeito de bolhas de ar sobre a percentagem da resistência residual do asfalto. Para finalidades do experimento, bolhas de ar são controladas em três níveis: baixo (2 – 4%), médio (4 – 6%) e alto (6 – 8%). Os dados são mostrados na seguinte tabela.

| Bolhas de Ar | Resistência Residual (%) | | | | | | | |
|--------------|--------------------------|----|-----|----|----|----|----|----|
| Baixa | 106 | 90 | 103 | 90 | 79 | 88 | 92 | 95 |
| Média | 80 | 69 | 94 | 91 | 70 | 83 | 87 | 83 |
| Alta | 78 | 80 | 62 | 69 | 76 | 85 | 69 | 85 |

(a) Os diferentes níveis de bolhas de ar afetam significativamente a resistência média retida? Use $\alpha = 0,01$;

Resolução:

A seguir, temos a formulação da hipótese e os procedimentos para o teste:

(i) Formulação da hipótese de interesse:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = 0 \\ H_1 : \text{caso contrário} \end{cases}$$

(ii) Fixar o nível de significância do teste e o valor F tabelado:

Usaremos $\alpha = 0,01$. Trabalharemos com uma F com 2 grau de liberdade no numerador e 21 graus de liberdade no denominador, logo pelo minitab o valor de F tabelado deve ser igual a:

Inverse Cumulative Distribution Function

F distribution with 2 DF in numerator and 21 DF in denominator

$$P(X \leq x) = 0,99 \quad x = 5,78042$$

Logo, temos: $F_{0,01;2;21} = 5,78042$.

(iii) *Determinação da região crítica e da região de aceitação:*

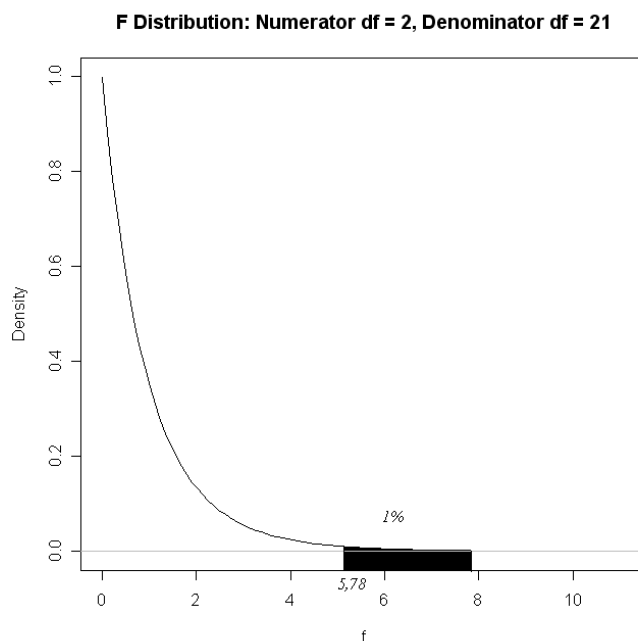


Figura 19: Determinação da região crítica.

(iv) *Cálculo do valor da Estatística de Teste é dado por:* $F_{cal} = \frac{R_c^2 - R_I^2}{\frac{nr}{(1-R_c^2)} \cdot \frac{1}{n-np}}$

Onde R_c^2 é definido como o coeficiente de determinação do modelo completo; R_I^2 é o coeficiente de determinação do modelo incompleto; nr é definido como o número de restrições e np é o número de parâmetros.

Logo, como $n = 24$; $R_c^2 = 0,442$; $nr = 2$ e $np = 3$, temos:

$$F_{cal} = \frac{\frac{0,442}{2}}{\frac{(1-0,442)}{24-3}} = 8,30$$

(v) *Conclusão: Portanto, ao nível de 1% de significância, existem fortes evidências amostrais que nos levam a rejeição da hipótese nula, ou seja, os diferentes tipos de bolhas afetam significativamente a resistência média retida ao nível de 1% de significância.*

- (b) Encontre o P para a estatística F calculada no item (a);

Resolução:

O p -valor é dado pelo complemento:

Cumulative Distribution Function

F distribution with 2 DF in numerator and 21 DF in denominator

$$x \quad P(X \leq x)$$

$$8,3 \quad 0,997793$$

Logo, p -valor: $1-0,997793=0,002207$.

- (c) Analise os resíduos provenientes desse experimento;

Resolução:

A seguir, temos os gráficos de resíduos:

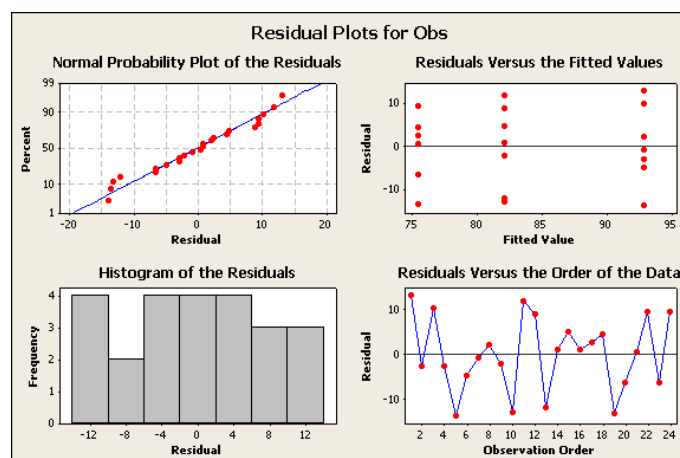


Figura 20: Gráfico de resíduo.

Logo, através da análise do gráfico de resíduos acima, é possível concluir que não há indícios de falta de adequação do modelo e das suposições, logo as conclusões por meio dos testes realizados podem ser consideradas válidas.

- (d) Encontre um intervalo de confiança de 95% para a resistência residual, em que há um nível alto de bolhas de ar;

Resolução:

Temos: $\bar{y}_1 = 75$ $t_{0,025;21} = 2,0796$ $MQ_E = 74,08$

$$\left[\bar{y}_1 \pm t_{0,025;21} \cdot \sqrt{\frac{MQ_E}{n}} \right] = [71,02; 79,98]$$

- (e) *Encontre um intervalo de confiança de 95% para a diferença na resistência média residual, nos níveis baixo e alto de bolhas de ar.*

Resolução:

$$\mu_i - \mu_j = (\bar{y}_i - \bar{y}_j) \pm t_{\alpha/2; k(n-1)} \sqrt{\frac{2 \cdot QM_{Res}}{n}}$$

Onde $t_{\alpha; k(n-1)}$ é o valor tabelado da distribuição *t-Student* com $k(n-1)$ graus de liberdade e $QM_{Res} = QM_{erro}$

Resolução:

$$\left[17,375 \pm t_{0,005;21} \cdot \sqrt{\frac{2 \cdot 74,08}{8}} \right] = [8,423; 26,327]$$

Exercício 10. No artigo intitulado “O efeito do Projeto do Bocal na Estabilidade e Desempenho de Jatos Turbulentos de Água” (The Effect of Nozzle Design on the Stability and Performance of Turbulent Water jets), na revista *Fire Safety Journal*, Vol. 4, agosto de 1981, C. Theobald descreve um experimento em que uma medida da forma foi determinada para vários tipos diferentes de bocais, com níveis diferentes de velocidade do jato de saída. O interesse nesse experimento está principalmente no tipo de bocal, sendo a velocidade um fator que provoca distúrbio. Os dados são apresentados a seguir:

| Tipo de Bocal | Velocidade do Jato de Saída (m/s) | | | | | |
|---------------|-----------------------------------|-------|-------|-------|-------|-------|
| | 11,73 | 14,37 | 16,59 | 20,43 | 23,46 | 28,74 |
| 1 | 0,78 | 0,80 | 0,81 | 0,75 | 0,77 | 0,78 |
| 2 | 0,85 | 0,85 | 0,92 | 0,86 | 0,81 | 0,83 |
| 3 | 0,93 | 0,92 | 0,95 | 0,89 | 0,89 | 0,83 |
| 4 | 1,14 | 0,97 | 0,98 | 0,88 | 0,86 | 0,83 |
| 5 | 0,97 | 0,86 | 0,78 | 0,76 | 0,76 | 0,75 |

(a) O tipo de bocal afeta a medida da forma? Compare os bocais, usando os diagramas de caixa e a análise de variância;

Resolução:

Analysis of Variance for Obs, using Adjusted SS for Tests

| Source | DF | Seq SS | Adj SS | Adj MS | F | P |
|--------|----|----------|----------|----------|------|-------|
| Tipo | 4 | 0,102180 | 0,102180 | 0,025545 | 8,92 | 0,000 |
| Vel | 5 | 0,062867 | 0,062867 | 0,012573 | 4,39 | 0,007 |
| Error | 20 | 0,057300 | 0,057300 | 0,002865 | | |
| Total | 29 | 0,222347 | | | | |

$S = 0,0535257$ $R-Sq = 74,23\%$ $R-Sq(adj) = 62,63\%$

Abaixo, temos os box-plot:

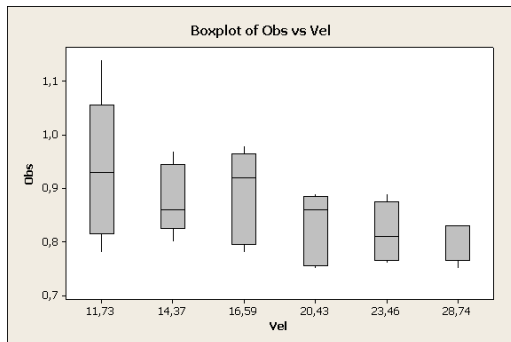


Figura 21: Veloc. x Observações

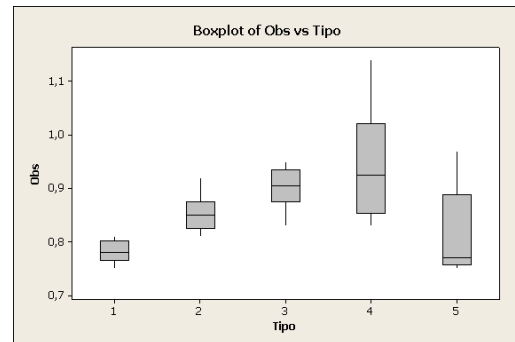


Figura 22: Tipo x Observações

Logo, podemos concluir pela análise de variância e pela análise dos diagramas de caixas, que o tipo de bocal afeta a medida da forma.

- (b) Use o método gráfico da Seção 12.3.1 para determinar diferenças específicas entre os bocais. Um gráfico da média (ou desvio-padrão) das medidas da forma versus o tipo de bocal ajuda nas conclusões?

Resolução:

Sim. As diferenças encontradas foram: 0.78; 0.81; 0.85; 0.90 e 0.94

- (c) Analise os resíduos desse experimento.

Resolução:

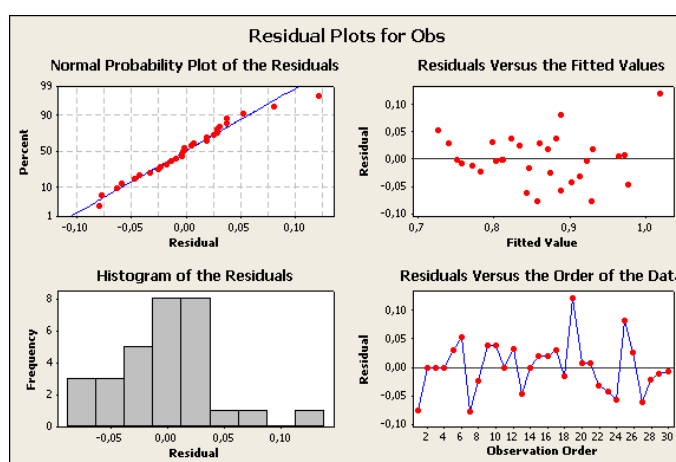


Figura 23: Resíduos

Observando o gráfico de resíduos contra a ordem de coleta dos dados não foi possível observar quaisquer configurações especiais que revelassem uma associação dos resíduos com a ordem de coleta das observações, o que indicou que a suposição de independência não parecia ter sido violada. O gráfico de resíduos contra médias nos mostra que as faixas de dispersão dos resíduos em cada média (tratamento) possuem comprimentos aproximadamente iguais, o que indica que a suposição de igualdade de variâncias poderia ser considerada válida e o gráfico de probabilidade normal nos mostra que os pontos estão distribuídos ao longo de uma reta, o que nos indica que os componentes de erro do modelo segue uma distribuição normal. Portanto, como não foram encontrados indícios de falta de adequação do modelo e das suposições a ele associadas, as conclusões obtidas por meio dos testes acima podem ser consideradas válidas.