Estatística

Prova 1 — MATD48 : Planejamento de Experimentos A

2021-1

Experimento de resistência a tensão: Um engenheiro de produção tem como interesse investigar a resistência a tensão de uma nova fibra sintética que utilizará na produção de camisas. Por experiência, o engenheiro sabe que a resistência é afetada pela proporção de algodão utilizado na fibra. Suspeita também que, com o aumento da proporção de algodão a resistência aumenta, pelo menos em um princípio; que a proporção de algodão deve variar entre 10 a 40 por cento no produto final. O engenheiro decide realizar um experimento considerando 5 níveis de proporção de algodão: 15, 20, 25, 30 e 35% e em cada nível produzir 5 exemplares. Trata-se de um experimento com um único fator com 5 níveis do fator e 5 réplicas. A Tabela a seguir mostra a distribuição aleatória do desenho inteiramente casualizado (DIC), assim como a resistência à tensão observada no experimento em $lb/pulg^2$.

Índice	Aleatorização	Porcentagem de algodão (%)	Tensão ($lb/pulg^2$)
1	8	20	12
2	18	30	19
3	10	20	17
4	23	35	7
5	17	30	25
6	5	15	7
7	14	25	14
8	6	20	12
9	15	25	18
10	20	30	22
11	9	20	18
12	4	15	7
13	12	25	18
14	7	20	18
15	1	15	15
16	24	35	10
17	21	35	11
18	11	25	19
19	2	15	11
20	13	25	19
21	22	35	15
22	16	30	19
23	25	35	11
24	19	30	23
25	3	15	9

Questões:

1. Suponha que temos **a** tratamentos ou níveis diferentes de um único fator que desejemos comparar. Consideremos y_{ij} a resposta observada em cada um dos **a** tratamentos como uma variável aleatória caracterizada pelo modelo

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}, \tag{1}$$

onde $\epsilon_{ij} \sim N(0, \sigma^2)$, para todo $i = 1, \ldots, a$ e $j = 1, \ldots, n$. Demostre que

$$\mathrm{E}\left[\frac{SQ_{\text{ERRO}}}{N-a}\right] = \sigma^2,\ N = a \times n.$$

2. Estimadores adequados para a média e efeito dos tratamentos do modelo (1) são dados por: $\widehat{\mu} = \overline{y}_{..}$ e $\widehat{\tau}_i = \overline{y}_{i.} - \overline{y}_{..}$, onde $\overline{y}_{..} = y_{..}/N$, $\overline{y}_{i.} = y_{i.}/n$, $y_{..} = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}$ e $y_{i.} = \sum_{j=1}^{n} y_{ij}$. Se consideramos a média do i-ésimo tratamento como $\mu_i = \mu + \tau_i$, para todo $i = 1, \ldots, a$ e σ^2 conhecido. Determine
   (a) Um intervalo de confiança de $100(1-\alpha)\%$ para a média μ_i do i-ésimo tratamento.
   (b) Um intervalo de confiança de $100(1-\alpha)\%$ para a diferencia de medias de dois tratamentos, por exemplo, $\mu_i - \mu_j$.

3. Suponha que no modelo (1) sejam observadas n_i réplicas para cada i-tratamento, e que $N = \sum_{i=1}^{a} n_i$. Demostre que

$$SQ_{\text{Tratamentos}} = \sum_{i=1}^{a} \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{N}.$$

4. Para o exemplo **experimento de resistência a tensão**.
   (a) Formule a hipóteses adequada para o experimento;
   (b) Defina e interprete o Erro Tipo I e II sobre as hipóteses em (a)
   (c) Construa a tabela ANOVA
   (d) A um nível de 5%, qual decisão o engenheiro irá tomar?.

   OBS: Aqui pode utilizar alguma livraria do R para confirmar seus cálculos algébricos.

5. Um procedimento comummente utilizado para testar

$$\begin{aligned} H_0 &: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_a^2 \\ H_1 &: \text{H}_0 \text{ não é verdade para algum } \sigma_i^2 \end{aligned}$$

é o teste Bartlett. Sob a hipótese nula, a estatística do teste é definido como $\chi^2 = 2,3026(q/c)$, onde

$$\begin{aligned} q &= (N-a)\log(S_p^2) - \sum_{i=1}^{a}(n_i-1)\log(S_i^2), \\ c &= 1 + \frac{1}{3(a-1)}\left(\sum_{i=1}^{a}(n_i-1)^{-1} - (N-a)^{-1}\right), \\ S_p^2 &= \frac{1}{N-a}\sum_{i=1}^{a}(n_i-1)S_i^2, \end{aligned}$$

onde S_i^2 representa a variância amostral da i-ésima população e N o número total de observações. Sem utilizar nenhuma libraria da linguagem de programação **R**, para os dados do **experimento de resistência a tensão** descrito ao início da prova, testar H_0 versus H_1 a um 5% de significância.

Prof. Jalmar Carrasco: carrascojalmar@gmail.com