O cimento de Portland

Universidade Federal da Paraíba - CCEN

Gabriel de Jesus Pereira

14 de março de 2024

O problema a ser resolvido e os dados

Exércicio 3.7.

The tensile strength of Portland cement is being studied. Four different mixing techniques can be used economically A completely randomized experiment was conducted and the following data were collected:

Mixing Technique	Tensile Strength (lb/in^2)			
1	3129	3000	2865	2890
2	3200	3300	2975	3150
3	2800	2900	2985	3050
4	2600	2700	2600	2765

O enunciado da questão nos diz que a resistência à tração do cimento de Portland está sendo estudada. Diz também que quatro diferentes técnicas de misturas podem ser utilizadas de forma econômica. Após isso, nos dá dados das quatro diferentes misturas da resistência à tração do cimento.

O que está sendo testado

O problema se trata de um experimento de fator único com $a = 4$ (quantidade de misturas diferentes) e $n = 4$ (réplicas de resistência à tração do cimento).

O que queremos testar é se as diferentes misturas afetam a resistência à tração do cimento.

Dessa forma, queremos testar se $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ contra a hipótese alternativa de que algumas médias são diferentes $\left(H_1 : \mu_i \neq \mu_j \mid \exists\ (i,\ j)\right)$. Assim, podemos interpretar H_0 como sendo a hipótese que nos diz que a técnica de mistura não afeta na resistência à tração do cimento. H_1 é a hipótese que nos diz que pelo menos uma das técnicas de mistura afeta na resistência à tração do cimento.

Suposições de normalidade e homogeneidade

A seguir, veremos os testes de hipótese e gráficos que sustentam as suposições necessárias para criação do modelo.

Normalidade

Tabela 2: Resultado dos testes para normalidade

Teste	Estatística	p-valor
Anderson-Darling	$A = 0.185$	0.890
Cramer-von Mises	$W = 0.028$	0.854
Lilliefors	$D = 0.118$	0.793
Shapiro-Francia	$W = 0.982$	0.947
Shapiro-Wilk	$W = 0.970$	0.846
Jarque-Bera	$JB = 0.688$	0.709

Mais da metade dos testes da Tabela 2 não rejeitam a hipótese de normalidade, e gráficamente pelo Q-Q plot da Figura 1 os resíduos parecem sim ter distribuição aproximadamente normal pois se assemelham aos quantis teóricos da distribuição normal.

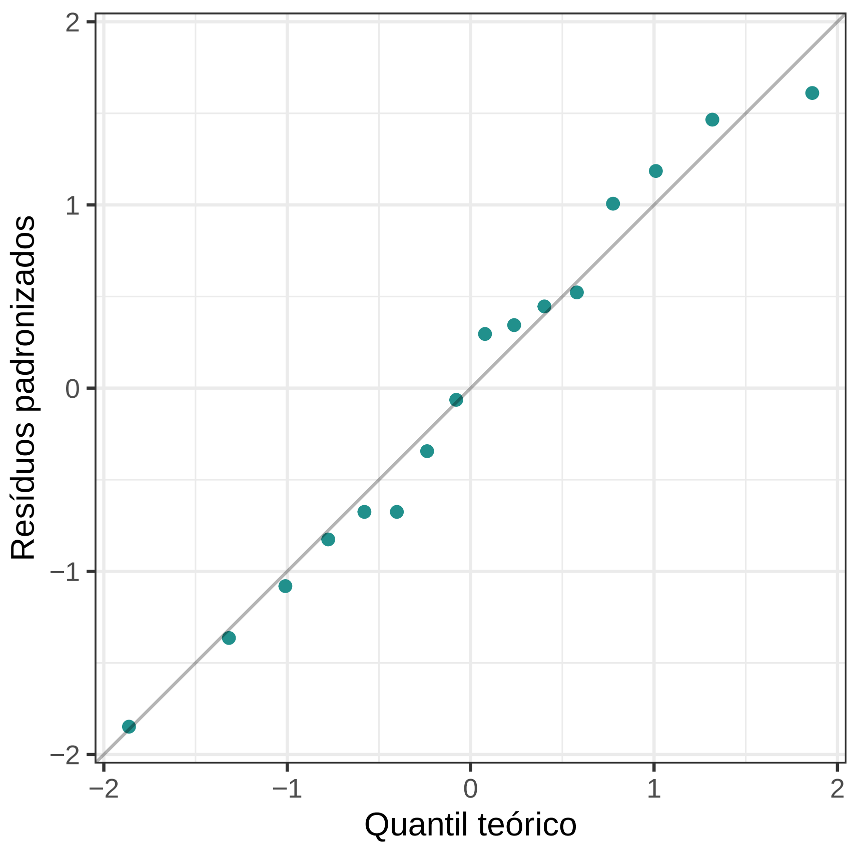
```r
# Dados
data <- tibble(
  `Técnica de mistura` = factor(rep(1:4, each = 4)),
  `Força de tensão` = c(3129, 3000, 2865, 2890,
                        3200, 3300, 2975, 3150,
                        2800, 2900, 2985, 3050,
                        2600, 2700, 2600, 2765)
)


# modelo One Way ANOVA
modelo <- aov(`Força de tensão` ~ `Técnica de mistura`, data = data)

# Código para criação de gráfico Q-Q plot
data |>
  ggplot(aes(sample = rstandard(modelo))) +
  geom_qq(color = "#21908C") +
  geom_abline(alpha = 0.3) +
  tune::coord_obs_pred() +
```

```
  theme_bw() +
  labs(
    x = "Quantil teórico",
    y = "Resíduos padronizados"
    )
```

Figura 1: Q-Q plot dos resíduos padronizados do modelo



Homogeneidade

Tabela 3: Resultado dos testes para homogeneidade

Teste	Estatística	p-valor
Barlett	$K = 0.711$	0.870
Levene	$F = 0.183$	0.905

Todos os testes da Tabela 3 não rejeitam a hipótese de que as variâncias são iguais entre os grupos, com p-valores acima de 5%. Isso indica que a variância do modelo é constante para todos os valores estimados da variável resposta.

A tabela ANOVA

Agora vamos testar as hipóteses e verificar se as técnicas de mistura afetam na resistência à tração do cimento de Portland.

O primeiro passo será calcular as somas dos quadrados dos erros e dos tratamentos, além de encontrar o grau de liberdade.

Soma total dos quadrados: $SS_T = \sum_{i=1}^{n} \sum_{i=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{N}$

Soma dos quadrados entre os tratamentos: $SS_{Tratamentos} = \frac{\sum_{i=1}^{n} y_{i.}^2}{n} - \frac{y_{..}^2}{N}$

Soma dos quadrados dentro dos tratamentos: $SS_E = SS_T - SS_{Tratamentos}$

Agora que temos as somas dos quadrados, podemos calcular os erros quadráticos médios, que podem ser calculados com as seguintes formas:

Erro quadrático médio entre os tratamentos: $MS_{Tratamentos} = \frac{SS_{Tratamentos}}{a-1}$

Erro quadrático médio dentro dos tratamentos: $MS_E = \frac{SS_E}{N-a}$

Em que $a - 1$ e $N - a$ são os graus de libertade entre e dentro dos tratamentos, respectivamente.

Dessa forma, chegamos na seguinte tabela ANOVA:

Tabela 4: Tabela da ANOVA

Variação	Soma dos Quadrados	Graus de Liberdade	Erro quadrático médio	F_0	$p-$ $valor$
$SS_{Tratamentos}$	489740.2	3	$MS_{Tratamentos} =$ 163246.73	12.73	4.89-4
SS_E	153908.2	12	$MS_E = 12825.69$		
SS_T	643648.4	15			

Dessa forma, vemos que $F_0\,(12.73) > F_{0.05,3,12}\,(3.49)$. Assim, rejeitamos H_0 e concluímos que algumas das diferentes técnicas de misturas afetam a resistência à tração do cimento.

Contrastes

```r
Contrastes <- tibble(
  C1 = c(1, -1, 0, 0),
  C2 = c(1, 1, -1, -1),
  C3 = c(0, 0, 1, -1)
  )
contrasts(data$`Técnica de mistura`) <- Contrastes |>
  as.matrix()

modelo_contraste <- aov(`Força de tensão` ~ `Técnica de mistura`, data = data)
```

	Coeficientes	Erro padrao	Estatistica t	$Pr(> \|t\|)$
Intercepto	2931.81	28.31	103.551	$< 2e-16$
$\mu_1 - \mu 2 = 0$	-92.63	40.04	-2.313	0.0392
$\mu_1 + \mu 2 - \mu_3 - \mu_4 = 0$	131.81	28.31	4.656	0.0005
$\mu_3 - \mu 4 = 0$	133.75	40.04	3.340	0.0058

Assim, concluímos analisando o p-valor que há diferenças significativas na resistência à tração entre os níveis 1 e 2 e entre o 3 e 4. Ainda, a média do nível 1 e 2 diferem significamente do média do nível 3 e 4.

Comparações múltiplas

Fisher Least Significant Difference (LSD)

$LSD = t_{\alpha/2,N-a}\sqrt{\frac{2MS_E}{n}} = 174.4797$

```r
LSD = LSD.test(modelo, "Técnica de mistura")

Diferenças <- data |>
  group_by(`Técnica de mistura`) |>
  summarise(Média = mean(`Força de tensão`)) |>
  (\(x) {
    n = nrow(x)
    all_dif <- sapply(1:(n - 1), FUN = \(i) {
      diffs <- sapply((i + 1):n, FUN = \(j) {
        x$Média[i] - x$Média[j]
      })
      c(rep(NA, i - 1), diffs)
    }) |>
      t() |>
      as_tibble()
    colnames(all_dif) <- 2:4
    all_dif
  })()
```

Tabela 6: Diferença das médias

	$\bar{y}_{2\cdot}$	$\bar{y}_{3\cdot}$	$\bar{y}_{4\cdot}$
$\bar{y}_{1\cdot}$	-185.25	37.25	304.75
$\bar{y}_{2\cdot}$		222.50	409.00
$\bar{y}_{3\cdot}$			267.50

Tabela 7: Significantemente diferente

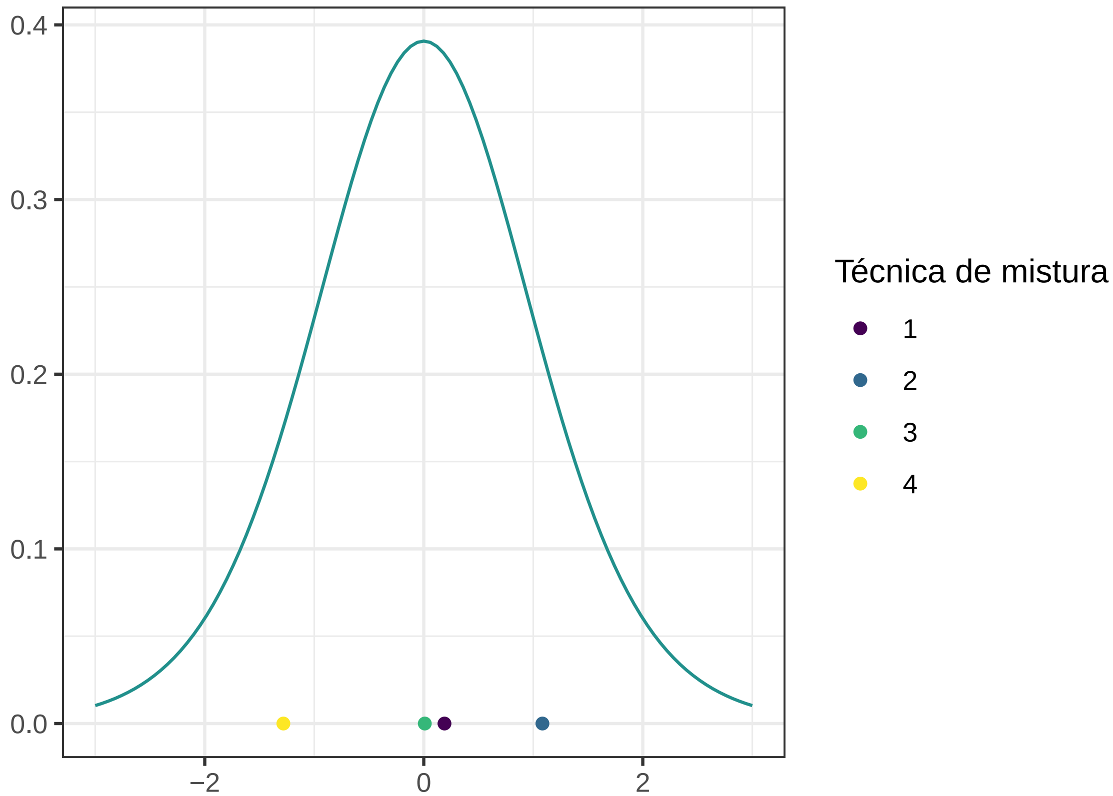
	$\bar{y}_{2\cdot}$	$\bar{y}_{3\cdot}$	$\bar{y}_{4\cdot}$
$\bar{y}_{1\cdot}$	1	0	1
$\bar{y}_{2\cdot}$		1	1
$\bar{y}_{3\cdot}$			1

Pelos resultados do Least Significant Difference (LSD), vemos que a única diferença que não é significantemente diferente é a diferença entre a média do grupo 1 e 3. Os outros resultados puderam ser comprovados também pelos contrastes, que vemos que há diferenças significativas na resistência à tração entre os grupos considerados.

Agora calculando: $S_{\bar{y}_{i.}} = \sqrt{\frac{MS_E}{n}} = 56.625$

```r
data |>
  mutate(`Força de tensão` = (`Força de tensão` - mean(`Força de tensão`)) / sd(`Força de tensã
  group_by(`Técnica de mistura`) |>
  summarise(media = mean(`Força de tensão`)) |>
  ggplot(aes(x = media, y = 0)) +
  xlim(-3, 3) +
  geom_function(fun = \(x) dt(x, df = 12), color = "#21908C") +
  geom_point(aes(color = `Técnica de mistura`)) +
  scale_color_viridis_d() +
  labs(x = "", y = "") +
  theme_bw()
```

Figura 2: Gráfico para comparar as médias da resistência à tração



Pelo gráfico da Figura 2, podemos concluir também sobre as diferenças. Vemos que a média da mistura 1 e mistura 3 estão bastante próximas, mostrando que sua diferenção não é significativa. Já para as outras diferenças, vemos que estão distantes, concluindo sobre as diferenças significativas que foram obtidas nos resultados anteriores.

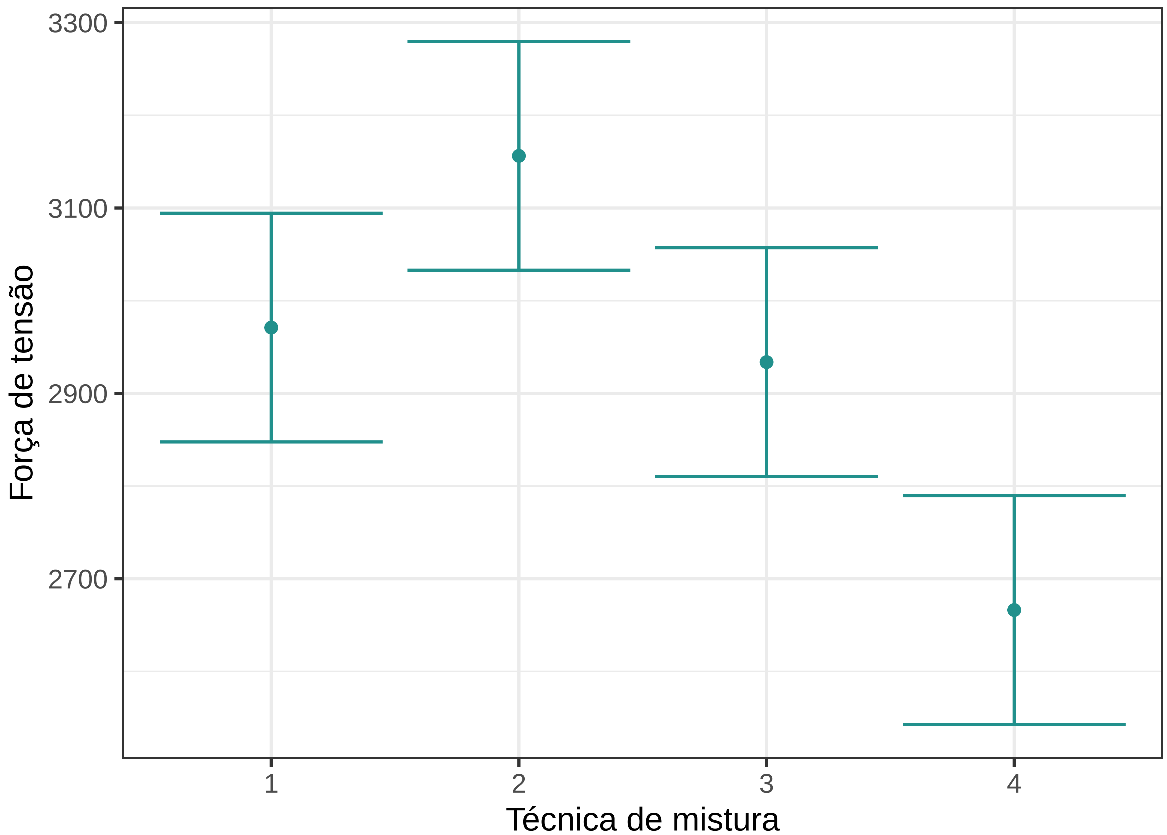
```
LSD$means |>
  select(`Força de tensão`, LCL, UCL) |>
  knitr::kable()
```

Figura 3: Tabela do intervalo de confiança de 95% de cada técnica de mistura

Força de tensão	LCL	UCL
2971.00	2847.624	3094.376
3156.25	3032.874	3279.626
2933.75	2810.374	3057.126
2666.25	2542.874	2789.626

```
LSD$means |>
  select(`Força de tensão`, LCL, UCL) |>
  mutate(`Técnica de mistura` = factor(1:4)) |>
  ggplot(aes(x = `Técnica de mistura`, y = `Força de tensão`)) +
  geom_point(color = "#21908C") +
  geom_errorbar(aes(ymin = LCL, ymax = UCL), color = "#21908C") +
  theme_bw()
```

Figura 4: Gráfico do intervalo de confiança para cada técnica de mistura



Os intervalos de confiança da técnica de mistura 1 e 3 estão bastante próximas, podendo significar, mais uma vez, que essas técnicas tem o mesmo efeito sobre a resistência à tração.