UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE MATEMÁTICA - DEPARTAMENTO DE ESTATÍSTICA
DISCIPLINA: MATF16 - MODELOS LINEARES GENERALIZADOS A
PROF: LEILA AMORIM

TÓPICO: ANÁLISE DE DADOS

LABORATÓRIO 2

1. Sementes geneticamente similares são aleatoriamente designadas para serem cultivadas em um ambiente nutricionalmente enriquecido (grupo tratado) ou em condições padrão (grupo controle) usando um experimento completamente aleatorizado. Após o tempo pré-determinado todas as plantas são colhidas, desidratadas e pesadas. Os resultados, expressos em gramas, para 20 plantas de cada grupo são mostrados na tabela a seguir (Dobson, pg 38):

Tabela I. Peso seco de plantas cultivadas sob duas condições.

Grupo Tratamento		Grupo Controle	
4.81	5.36	4.17	4.66
4.17	3.48	3.05	5.58
4.41	4.69	5.18	3.66
3.59	4.44	4.01	4.50
5.87	4.89	6.11	3.90
3.83	4.71	4.10	4.61
6.03	5.48	5.17	5.62
4.98	4.32	3.57	4.53
4.90	5.15	5.33	6.05
5.75	6.34	5.59	5.14

É de interesse testar se há alguma diferença na produção entre os dois grupos. Seja Y_{jk} a k-ésima observação do j-ésimo grupo, onde $j = 1$ para o grupo de tratamento e $j = 2$ para o grupo controle, com $k = 1, 2, ..., 20$ para ambos os grupos. Assume-se que Y_{jk}'s são variáveis aleatórias independentes com $Y_{jk} \sim N(\mu_j, \sigma^2)$. A hipótese nula é $H_o : \mu_1 = \mu_2 = \mu$, ou seja, não há diferença entre os grupos, versus a hipótese alternativa $H_o : \mu_1 \neq \mu_2$.

a) Realize análise exploratória desses dados verificando as distribuições para cada grupo (análise gráfica) e calculando estatísticas sumárias. O que você pode concluir com base nestas avaliações?

b) Utilize um teste t-Student não pareado e calcule o intervalo de 95% de confiança para a diferença entre as médias desses grupos. Interprete esses resultados.

c) Os seguintes modelos podem ser usados para testar a hipótese nula contra a alternativa, onde: $H_o : E(Y_{jk}) = \mu$; $Y_{jk} \sim N(\mu, \sigma^2)$
$H_1 : E(Y_{jk}) = \mu_j$; $Y_{jk} \sim N(\mu_j, \sigma^2)$
para $j = 1, 2$ e $k = 1, 2, ..., 20$. Encontre as estimativas de máxima verossimilhança para os parâmetros μ, μ_1 e μ_2. Qual sua conclusão a respeito dessas hipóteses?

d) Calcule os resíduos para o modelo definido em H_1 e use-os para explorar as suposições distribucionais do modelo.

2. Os pesos, em quilogramas, de vinte homens antes e após a participação em um programa de perda de peso são mostrados na tabela a seguir (Egger et al., 1999). O interesse é saber se, em média, há uma retenção de perda de peso doze meses após a participação no programa.

Tabela II. Peso antes e após participação em um programa de perda de peso.

Registro	Antes	Após	Registro	Antes	Após
1	100.8	97.0	11	105.0	105.0
2	102.0	107.5	12	85.0	82.4
3	105.9	97.0	13	107.2	98.2
4	108.0	108.0	14	80.0	83.6
5	92.0	84.0	15	115.1	115.0
6	116.7	111.5	16	103.5	103.0
7	110.2	102.5	17	82.0	80.0
8	135.0	127.5	18	101.5	101.5
9	123.5	118.5	19	103.5	102.6
10	95.0	94.2	20	93.0	93.0

Seja Y_{jk} o peso do k-ésimo homem no j-ésimo tempo, onde j=1 antes do programa e j=2 doze meses depois. Assuma que Y_{jk}'s são variáveis aleatórias independentes com $Y_{jk} \sim N(\mu_j, \sigma^2)$ para $j = 1, 2$ e $k = 1, 2, ..., 20$.

a) Use o teste t não pareado para testar as hipóteses $H_o : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$.

b) Seja $D_k = Y_{1k} - Y_{2k}$, para $k = 1, 2, .., 20$. Formule modelos para testar H_o versus H_1 usando os D_k's.

c) Use o teste t-pareado para testar as hipóteses em (a). As conclusões são as mesmas em todas as análises realizadas?