1. (Neter et el., 1996, p. 449). No arquivo vendas.dat são descritas informações a respeito das vendas no ano anterior de um tipo de telhado de madeira em 26 filiais de uma rede de lojas de construção. As variáveis estão colocadas na seguinte ordem: (i) telhados, total de telhados vendidos (em mil metros quadrados), (ii) gastos, gastos pela loja com promoções do produto (em mil USD), (iii) clientes, número de clientes cadastrados na loja (em milhares), (iv) marcas, número de marcas concorrentes do produto e (v) potencial, potencial da loja (quanto maior o valor maior o potencial). Um dos objetivos do estudo com esse conjunto de dados é tentar prever o número esperado de telhados vendidos dadas as variáveis explicativas. Faça inicialmente uma análise descritiva construindo, por exemplo, os diagramas de dispersão de cada variável explicativa contra a variável resposta telhados. Calcule também as correlações entre as variáveis. Use os métodos stepwise e AIC para selecionar um modelo de regressão normal linear. Se o modelo selecionado for diferente pelos dois métodos, adote algum critério para escolher um dos modelos. Interprete os coeficientes estimados do modelo selecionado. Faça uma análise de diagnóstico para verificar se existem afastamentos sérios das suposições feitas para o modelo e se existem observações discrepantes.

Dados: https://www.ime.usp.br/~giapaula/vendas.dat

1. A distribuição binomial B(m, θ), com 0 < θ < 1 e m, o número conhecido de ensaios independentes, é usada para análise de dados na forma de proporções e tem função de probabilidade:



Coloque na forma:



1. A distribuição de Poisson P(θ) de parâmetro θ > 0, usada para analisar dados na forma de contagens, tem função de probabilidade:



Coloque na forma:



4. A distribuição binomial B(m, θ), com 0 < θ < 1 e m, o número conhecido de ensaios independentes, é usada para análise de dados na forma de proporções e tem função de probabilidade:

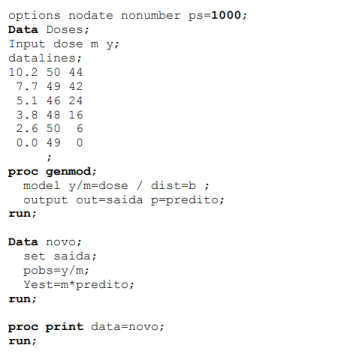


Coloque na forma:

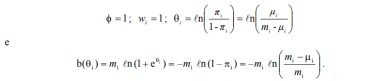


5. Se Y tem distribuição exponencial de média unitária, mostre que a função geratriz de momentos de Y = log(X) é igual a M(t) = Γ(1 + t) e que a sua f.d.p. é f(y) = exp(y − ey).

6. Análise o modelo abaixo:



7. Sejam Yi variáveis aleatórias representando contagens de sucessos em amostras independentes de tamanhos mi. Supondo que Yi ~ Bin(mi, Πi) , então,

Encontre Sp.

8. Seja Y1, Y2, ..., Yn (grande) uma amostra aleatória de uma distribuição N(µ,σ2) com µ desconhecido e σ2 conhecido. Visto como modelo linear generalizado tem-se:

i) somente um parâmetro de interesse, µ ;

ii) não há variáveis explicativas e,

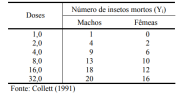
iii) a função de ligação é a identidade, isto é, η = µ .

O logaritmo da função de verossimilhança é

Encontre:

 ; E(U) ; ; estatística de escore (E) e sua distribuição.

9. Amostras de 20 insetos, Heliothis virescens (praga do algodão), resistentes a cypermethrin, foram expostas a doses crescentes do inseticida, dois dias depois da emergência da pupa. Após 72h foram contados os números de insetos mortos e os resultados obtidos estão abaixo.



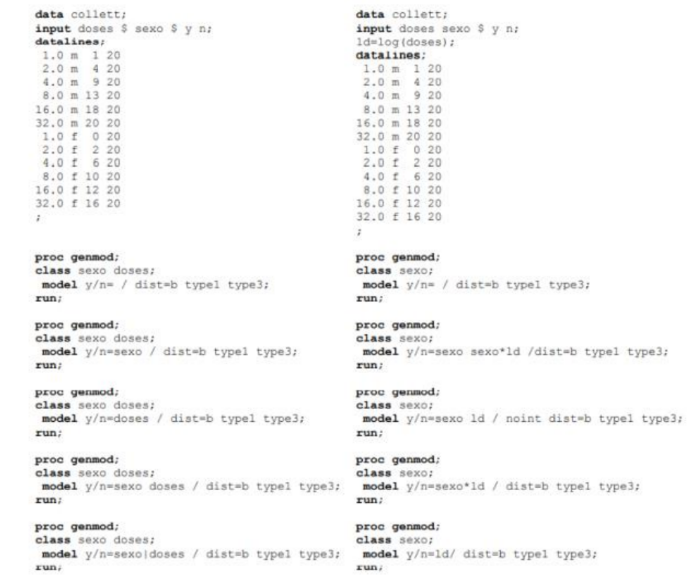
Considerações:

- variável resposta: Yi = número de insetos mortos em amostras de tamanho mi = 20; - distribuição: Binomial;

- parte sistemática: modelos de regressão;

- objetivo: determinação de doses letais para machos e fêmeas e verificação se são diferentes. Compare os modelos abaixo.

Segue o programa:



10.

