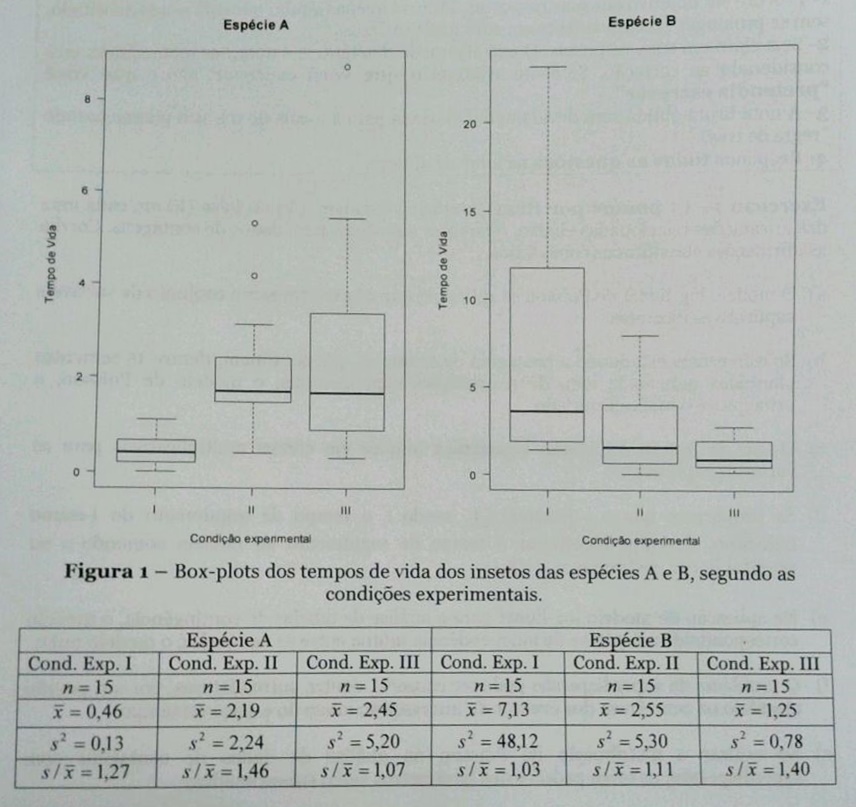
**PROVA 2 – MODELOS LINEARES GENERALIZADOS**

**Exercício 1 – (5 pontos por item)** Assinale verdadeiro (V) ou falso (F) em cada uma das afirmações relacionadas abaixo, referentes a modelos para dados de contagens. Corrija as afirmações classificadas como falsas.

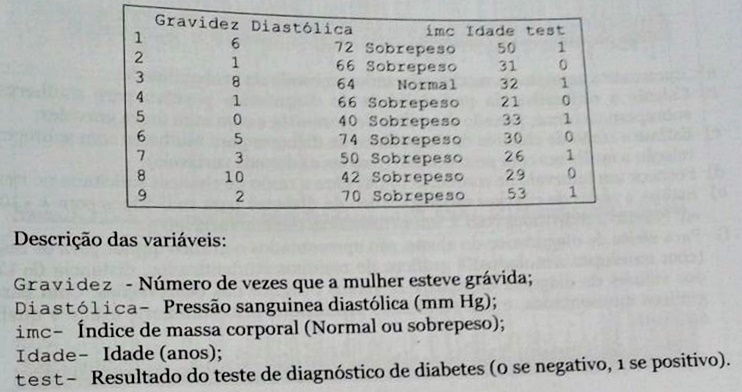
1. O modelo log-linear de Poisson se aplica sempre que tivermos um conjunto de variáveis explicativas discretas;
2. Se estivermos estudando a contagem de sementes que germinam, dentre 15 sementes plantadas para cada uma de 5 condições experimentais, o modelo de Poisson, a princípio, é o mais apropriado.
3. O uso da função de ligação logarítmica implica em efeitos multiplicativos para as variáveis explicativas;
4. Se propormos *yi ~ Poisson (ti λi)*, sendo ti o tempo de seguimento do i-ésimo indivíduo, devemos incorporar o tempo de seguimento ao modelo somando-o ao preditor linear, na forma: ni = **xi‘ β** + ti ;
5. Na aplicação do modelo log-linear para a análise de tabelas de contingência, o modelo correspondente à hipótese de independência mútua entre as variáveis é o modelo nulo;
6. O problema da superdisposição pode ser causado, dentre outros fatores, por um padrão aleatório na ocorrência dos eventos de interesse ao longo do espaço ou tempo;
7. Ao utilizar a distribuição de Poisson na análise de dados de contagens com superdispersão, os erros padrões dos parâmetros serão subestimados;
8. Ao usar o modelo de regressão quase-Poisson, considerando *V(µi)= ϕ µ,* com *ϕ* a ser estimado, as estimativas pontuais dos *β’s* serão idênticas às produzidas pelo modelo de regressão Poisson.

**Exercício 2 – (40 pontos)** Um experimento tem como objetivo comparar os tempos médios de vida de duas espécies de insetos (A e B) submetidos a três condições experimentais distintas (I, II e III). Para isso, foram observados os tempos de vida de 45 insetos de cada espécie, com 15 insetos de cada espécie submetidos a cada condição experimental. Na sequência são apresentados gráficos e algumas medidas descritivas baseadas nos resultados do experimento:

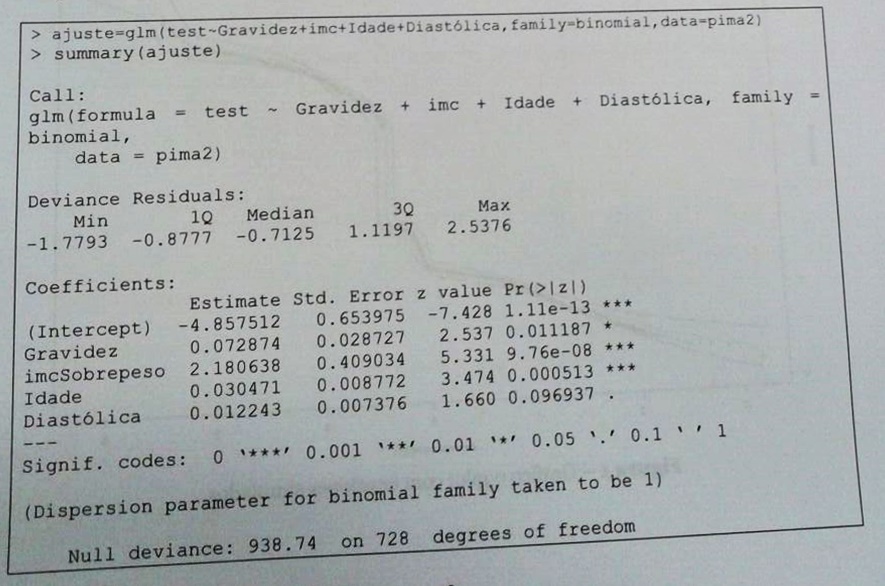


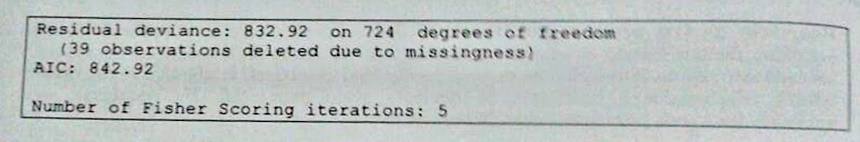
Para o problema apresentado, proponha um MLG em duas etapas, conforme visto em aula, especificando, num primeiro momento, a distribuição da resposta condicional às covariáveis e, posteriormente, a relação entre a distribuição da resposta e o preditor linear. Não se esqueça de deixar claro quem são as variáveis resposta e explicativas e como são inseridas no modelo. Justifique suas especificações.

**Exercício 3 - (10 pontos por item)** Os dados apresentados na sequencia foram extraídos de um estudo conduzido pelo Instituto de Diabetes e Doenças Digestivas, baseado em 729 mulheres adultas de uma comunidade indígena. O objetivo é identificar fatores relacionados à incidência de diabetes nessa população. Na sequência são apresentadas as seis primeiras linhas da base de dados:



Para a análise desses dados, foi ajustado um modelo de regressão logística, considerando o resultado do teste como variável resposta e as demais variáveis como explicativas. Nenhuma interação foi incluída ao modelo. O quadro apresentado na sequencia contém o resumo do modelo ajustado:





Com base no modelo ajustado:

1. Apresente a equação do modelo ajustado, na escala de probabilidade;
2. Calcule a estimativa da probabilidade de diagnóstico positivo para mulheres com sobrepeso, 40 anos, pressão diastólica de 50mmHg e com uma única gravidez;
3. Estime a razão de chances de diagnóstico de diabetes para mulheres com sobrepesp em relação a mulheres com peso normal (fixadas as demais variáveis);
4. Forneça um intervalo de confiança (95%) para a razão de chances solicitada no item b;
5. Estime a razão de chances de diagnóstico de diabetes para indivíduos com *k* + 10 anos em relação a indivíduos com *k* anos (fixadas as demais variáveis);
6. Para efeito de diagnóstico do ajuste, são apresentados o gráfico qqplot para resíduos (com envelopes simulados) e gráficos de resíduos studentizados, distância de Cook e dos valores da diagonal da matriz *H* versus o índice das observações. Com base nos gráficos apresentados, e nos resultados produzidos pelo summary, avalie a qualidade do ajuste.

