

**Departamento de Estatística – IMECC – UNICAMP**  
**ME 323 A - Introdução aos Modelos Probabilísticos - Prof. Amorim**  
**Quinta lista de exercícios – devolução: 19/06/2020**

Nome: \_\_\_\_\_

	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	g
RA							

1. Uma grade ortogonal, como sugere a figura ao lado, é formada por  $a$  linhas e  $b$  colunas; ela contém, portanto,  $a.b$  células. Existem dois tipos de célula nesta grade, representados por 0 e 1, e a fração de células do tipo 1, que chamaremos de  $p$ , é desconhecida. Como  $m$  e  $n$  são muito grandes, e a análise de uma célula para determinar o seu tipo é cara e demorada, a inspeção de todas as células é totalmente fora de consideração. No entanto, vamos admitir, é muito importante obter uma estimativa precisa de  $p$  e você é desafiado a obter esta estimativa através da aplicação da Teoria da Probabilidade. Para isto você selecionará, por sorteio aleatório simples, com reposição (isto é, uma célula sorteada continua participando dos sorteios seguintes, podendo ser sorteada novamente, de forma que a cada sorteio é feito sobre o mesmo conjunto de  $a.b$  células da grade).



Seja  $a=(27 + c_6 + g) \times 10^5$  e  $b=(5 + c_5 - g) \times 10^3$

- Como você selecionaria uma a.a.s.-c.r. de  $n$  células da grade?
  - Considere uma amostra piloto de  $n=200+20c_6+10g$  células, com  $x=...$  casos positivos. Estime  $p$
  - determine o tamanho amostral que garanta uma estimativa de  $p$  com erro absoluto inferior a um ponto percentual, com probabilidade pelo menos 0,95.
  - E com erro relativo inferior a 1%?
2. Imagine agora que as linhas da grade acima são  $a$  estudantes, e que as colunas são as  $b$  questões de um super teste, cobrindo de forma homogênea e completa todo o programa oficial de, digamos, Matemática do Ensino Fundamental. Os valores de cada célula se referem à qualidade da resposta que o aluno daria à questão correspondente: 0 para errada e 1 para correta.
- Explique o significado de  $p$ .
  - Imagine agora que o valor  $X$  associado a cada célula possa ser qualquer valor no intervalo  $[0, 1]$ , referindo-se a possíveis soluções parciais do aluno à questão. Interprete  $\bar{X}$ .
  - Determine um limitante superior para  $V(\bar{X})$ .
  - Com base na resposta acima, defina o tamanho amostral  $n_0$  para uma amostra piloto que estime  $\bar{X}$  com erro inferior a 0,05 com probabilidade 0,95.
  - Suponha a seguinte amostra e os seguintes resultados:

$$n_0=100, \sum_{i=1}^{100} x_i = 30,32 \text{ e } \sum_{i=1}^{100} x_i^2 = 11,68$$

Construa um i95%c para  $\bar{X}$

- determine o tamanho amostral para garantir um i95%c com apenas dois pontos percentuais de largura.

3. Trabalhe sobre o conjunto de dados disponibilizado no Moodle, com N de alunos e M questões, para explorar algumas simulações.
  - a. Defina um plano amostral e estime a nota média, com um i95%c adequadamente preciso.
  - b. Nessa amostra ocorreram alunos do subgrupo 1 e do subgrupo 2. Verifique se existe diferença significativa de performance entre os dois subgrupos.
  - c. Analogamente, a amostra envolveu questões de dois subgrupos temáticos, A e B. Verifique se existe diferença significativa entre os níveis de desempenho dos alunos nesses dois grupos temáticos.
  - d. Construa uma tabela cruzada que sumarie simultaneamente, os resultados nos 4 subconjuntos implícitos.
  - e. Como você reestruturaria esse projeto.