\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Lista 1 – Experimento de Monte Carlo**

**Econometria -- EM 2021.02**

**Prof. Rudi Rocha**

**FGV EAESP**

**Nome dos integrantes do grupo:**

**Preencher aqui nomes:**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Instruções**

1. Este é um trabalho a ser realizado em duplas.
2. Cada dupla deverá trabalhar com a base de dados em Excel “lista1.xls”, disponível no e-class.
3. Cada dupla deverá responder aos itens perguntados neste mesmo arquivo em doc. Ao final, o arquivo deverá ser convertido e entregue em pdf, respeitando a formatação do doc.
4. A entrega da lista ocorrerá via e-class, em horário a ser definido.

\*\*\*

Nesta lista vamos realizar um experimento de Monte Carlo utilizando o MS Excel. A planilha sobre a qual vocês trabalharão é a “lista1.xlsx”, guia “Experimento MC”.

Nesta guia, vocês encontrarão na coluna A o identificador de 1000 observações. A coluna B gera um número aleatório entre 0 e 1 para cada uma dessas observações (usou-se a função ALEATÓRIO( ) do Excel), enquanto a coluna C retorna para cada um desses números aleatórios o inverso da função de distribuição acumulada da normal padrão, ou seja, uma Normal(0,1). Não se assustem com esse começo, essas primeiras colunas servem apenas para conseguirmos gerar uma variável aleatória para cada *i*.

Em seguida, geramos arbitrariamente uma coluna com a variável X, que assume um valor fixo para cada uma das 1000 observações (esta coluna está dada e foi definida previamente). Por fim, geraremos a coluna com a variável Y da seguinte forma. Em primeiro lugar, vamos escolher um parâmetro . Depois, você definirá o parâmetro β como sendo igual aos dois primeiros números de sua matrícula (de um/a dos/as integrantes da dupla), dividido por 1000. Ou seja, se a matrícula é composta pelos números 098100358, . Finalmente, você já será capaz de criar a coluna Y como:

Está definido então o nosso modelo linear populacional (afinal, fomos nós que definimos os parâmetros populacionais). Note também que, para cada realização de valores para a variável estocástica *u* (coluna C), temos duplas (*Yi, Xi*) distintas. Pensem nisso como a realização de uma amostra específica. Assim que abrirem a guia “Experimento MC”, vocês encontrarão um exemplo de como pode ser gerada a coluna Y. Já podemos começar o exercício.

**Parte 1**

**Questão 1** – Após gerar as suas próprias colunas de A a E, você estará em condições de estimar os parâmetros populacionais para a amostra aleatória de dados que você tem em mãos (colunas X e Y).

1. Utilize as fórmulas vistas em aula para os estimadores de MQO e estime e . Reporte os seus resultados e interprete as estimativas encontradas.

**Resposta:**

1. Você encontrou estimativas iguais aos parâmetros populacionais? Explique porque sim (ou porque não) e justifique a sua resposta.

**Resposta:**

1. Faça um gráfico de dispersão Y (eixo y) vs X (eixo x). Plote o seu gráfico abaixo e plote também a equação linear estimada (use o recurso do próprio Excel para estimar e plotar a sua equação).

**Resposta:**

1. Exercício extra: calcule também o SQE, o SQT, o SQR e o R2 desta regressão e reporte as estimativas abaixo.

**Resposta:**

**Obs**: Observação: você notará que a cada manipulação nos dados da planilha, o Excel gerará novos números aleatórios e, assim, novos dados para u e Y. Para realizar o exercício e responder aos itens acima, basta fixar a coluna de erros (copiar e colar como valores, por exemplo).

**Questão 2** – Vamos começar agora a replicar esse experimento. Em primeiro lugar, gere 20 novas colunas B e, para cada uma, gere novas colunas C e E mantendo-se fixa a coluna D e os parâmetros populacionais escolhidos na primeira questão.

Note que, para cada vez que você salvar o seu arquivo, o Excel atualizará automaticamente os números gerados aleatoriamente. O que fizemos aqui foi criar 20 novas amostras aleatórias para a nossa população de interesse. Se você deixar as fórmulas montadas, os exercícios abaixo poderão ser realizados rapidamente.

1. Para cada uma das 20 amostras geradas, calcule . Você terá gerado 20 muito provavelmente!) diferentes. Faça um gráfico/histograma contendo a frequência dessas estimativas e reporte abaixo. Responda: algum dos 20 estimados é idêntico ao parâmetro populacional ?

**Resposta:**

1. Calcule a média e a variância dos estimados utilizando as suas 20 estimativas. Reporte abaixo essas duas estimativas.

**Resposta:**

1. Exercício extra: se você conseguir trabalhar em algum software estatístico mais avançado, refaça os itens (a) e (b) com base em 100 repetições amostrais e depois com base em 1000 repetições amostrais: plote os gráficos e reporte as médias e variâncias dos estimados de em cada um dos dois casos. Responda: a média dos estimados está se aproximando do verdadeiro parâmetro populacional conforme aumenta o número de repetições?

**Resposta:**

**Questão 3** – Agora vamos gerar um problema: uma dependência entre u e X. Vamos trabalhar com a guia “VVO”. Nesta guia, vamos reconstruir a coluna C (dos u) da mesma forma como fazíamos anteriormente, mas agora somando um novo termo: 0,35\*X. Ou seja, ao invés de utilizarmos u no momento de gerar Y, utilizaremos u+0,35X. A partir de agora, repita exatamente os mesmos passos da Questão 2 e gere 20 novas amostras aleatórias para a população de interesse. Responda aos itens:

1. Para cada uma das 20 amostras geradas, calcule . Você terá gerado 20 : faça um gráfico/histograma contendo a frequência dessas estimativas e reporte abaixo. Responda: algum dos 20 estimados é idêntico ao parâmetro populacional ?

**Resposta:**

1. Calcule a média e a variância dos estimados utilizando as suas 20 estimativas. Reporte abaixo essas duas estimativas.

**Resposta:**

1. Exercício extra: se você conseguir trabalhar em algum software estatístico mais avançado, refaça os itens (a) e (b) com base em 100 repetições amostrais e depois com base em 1000 repetições amostrais: plote os gráficos e reporte as médias e variâncias dos estimados de em cada um dos dois casos. Responda: a média dos estimados está se aproximando do verdadeiro parâmetro populacional conforme o número de repetições aumenta?

**Resposta:**

1. Por fim, e o mais importante: você diria que a média de calculada com base nas suas novas 20 estimativas está mais próxima do verdadeiro parâmetro que a estimativa de encontrada na Questão 2 acima? O que deu errado, afinal?

**Resposta:**

\*\*\*

**Parte 2**

**Questão 4** – Retorne para a Questão 1 da primeira parte. Vamos realizar exercícios de inferência.

1. Estime o erro-padrão do seu estimador e obtenha um intervalo de confiança de 95% para o parâmetro populacional . Agora obtenha um intervalo de 99%.

**Resposta:**

1. Verifique se o parâmetro populacional que você definiu pertence aos intervalos obtidos. Lembre-se de que, na prática, nunca sabemos qual é o verdadeiro valor desse parâmetro – e é exatamente isso que queremos descobrir. Neste exercício prático nós o conhecemos porque, afinal, fomos nós mesmos que o criamos.

**Resposta:**

**Questão 5** – Vamos replicar novamente o experimento. Gere as 20 novas colunas B e, para cada uma, gere novas colunas C e E mantendo-se fixa a coluna D e os parâmetros populacionais escolhidos na primeira questão.

1. Para cada nova amostra: (i) calcule , (ii) calcule o seu respectivo erro padrão, (iii) obtenha um IC de 95% para o parâmetro populacional .

**Resposta:**

1. Utilizando as 20 amostras geradas na questão anterior, e os 20 intervalos de confiança estimados, responda: (i) quantos destes 20 intervalos contêm o verdadeiro parâmetro populacional (ii) com base na resposta anterior, você rejeitaria a hipótese nula de que é igual ao verdadeiro valor do parâmetro populacional que nós criamos? (iii) você rejeitaria a hipótese de que ?

**Resposta:**

1. Extra: repita os exercícios com 100 e 1000 repetições.

**Resposta:**

\*\*\*