Estatística e Probabilidade - Avaliação “Presencial” 02 - 2022/02

Prof. Hugo Carvalho

15/11/2022

– INSTRUÇÕES – LEIAM ATENTAMENTE! –

- A data limite de entrega da avaliação é sexta-feira 16/12/2022 às 12h. Avaliações entregues após esse prazo serão desconsideradas.
- A entrega deve ser feita exclusivamente através do Google Classroom, clicando na caixa “+ Adicionar ou Criar” dentro da postagem dessa lista, para então anexar um arquivo com sua resolução. Após isso, clique em “Entregar” para enviar sua resolução.
  Atenção: Somente anexar a resolução não é suficiente! O envio deve ser feito para que sua resolução de fato seja entregue.
- Você tem a liberdade de escrever sua resolução no computador (usando Word, LaTeX, dentre outros), ou manuscrito e depois escanear ou fotografar a sua resolução. Nesse último caso, tome cuidado para que o documento fique legível. No caso de fotografar, opte por utilizar luz natural e tome cuidado com sombras.
- **Dica**: Se for fotografar sua resolução com um *smartphone* ou *tablet*, utilize o aplicativo próprio da câmera, e evite fotografar através de WhatsApp, Telegram, Messenger, e outros. Os aplicativos de comunicação, ao utilizarem a câmera, fazem uma severa compressão da imagem, incorrendo em uma grande diminuição de sua qualidade. Para transferir a imagem do celular para o computador prefira fazer o envio por e-mail, ou acessando sua galeria de fotos através do Google Photos no computador (caso já utilize esse aplicativo para gerenciar suas fotos no aparelho).
- Independente do modo de escrita, a resolução deve ser entregue em um único documento, no formato PDF, com a resolução em pé (formato retrato). O *layout* da resolução não será levado em consideração na avaliação, porém o texto deve estar legível para ser corrigido.
  Atenção: Resoluções ilegíveis ou fora desse formato não serão corrigidas e serão desconsideradas.
- A troca de conhecimento na realização da avaliação é permitida e encorajada: ciência se faz com colaboração, e devemos seguir esse espírito aqui. Porém, cada aluno deverá ter a sua própria resolução, e cópias ou outras ilegalidades serão severamente punidas com a anulação da avaliação para o(s) aluno(s) suspeito(s).
- Todos os passos de sua resolução devem ser devidamente justificados.
- Ao entregar essa avaliação, você afirma ter lido e estar de acordo com essas regras, comprometendo-se a cumpri-las.

Questão 1: Sejam X e Y variáveis aleatórias independentes com distribuição de Poisson de parâmetros λ_1 e λ_2, respectivamente.

a) Mostre que $X+Y$ também tem distribuição de Poisson, mas com parâmetro $\lambda_1+\lambda_2$. ***(1,5)***
Dica: Seu objetivo é encontrar $\mathbb{P}(X+Y=z)$, para $z=0,1,2,3,\ldots$. Para isso, argumente que o evento $\{X+Y=z\}$ é igual ao evento $\bigcup_{x=0}^{z}\{X=x \text{ e } Y=z-x\}$, e use isso para continuar as contas. O fato de certos eventos serem disjuntos e o fato de X e Y serem independentes irá ser necessário daqui em diante.

b) Mostre que a distribuição condicional de $X|(X+Y=z)$ é uma Binomial com parâmetros z e $\lambda_1/(\lambda_1+\lambda_2)$. Se você não conseguiu fazer o item a), pode usar livremente seu resultado aqui. ***(2,0)***
Dica: Agora, você quer encontrar $\mathbb{P}(X=x|X+Y=z)$. Primeiramente, argumente porque x só pode tomar valores entre 0 e z. Para calcular a probabilidade desejada use a definição de probabilidade condicional e, em momento oportuno, argumente que o evento $\{X=x \text{ e } X+Y=z\}$ é igual ao evento $\{X=x \text{ e } Y=z-x\}$. Daí em diante, as contas sairão naturalmente.

c) Explique, intuitivamente, os resultados dos itens a) e b) acima, com base em interpretações das distribuições de probabilidade encontradas. ***(1,0)***
Dica: Crie um cenário fictício, onde alguma coisa é modelada por X e outra coisa é modelada por Y. Dentro desse cenário, argumente intuitivamente que é razoável que a distribuição de $X+Y$ seja também uma Poisson, e que também é razoável termos $X|(X+Y)$ como uma Binomial.

Questão 2: Dado a um número real, denote por $[a]$ a sua *parte inteira*, ou seja, $[a]$ é o maior inteiro que é menor ou igual à a. Seja $X \sim \text{Exp}(\lambda)$, e faça o que se pede abaixo:

a) Encontre a distribuição de probabilidade de $[X]$. ***(1,5)***
Dica: Primeiramente, argumente que, enquanto X assume valores reais positivos, a variável aleatória $[X]$ assume somente os valores $0,1,2,3,\ldots$. Após isso, seu objetivo será calcular $\mathbb{P}([X]=x)$, para $x=0,1,2,3,\ldots$. Para isso, será necessário que você relacione o evento $\{[X]=x\}$ com algum evento que diga respeito à variável aleatória X somente, pois você conseguirá calcular a sua probabilidade e chegar no resultado pretendido.

b) **(QUESTÃO BÔNUS)** Interprete o resultado obtido acima. Para isso, relacione esse resultado com alguma distribuição de probabilidade que já deu as caras em nosso curso, e faça um paralelo com a sua respectiva interpretação. ***(1,0)***

Questão 3: Seja X uma variável aleatória codificando uma característica de uma população cuja média μ é um valor desconhecido e queremos estimar. Por estudos anteriores, a variância de X é conhecida e denotada por σ^2. Para se estimar o valor de μ, coleta-se então dados, representados na forma de variáveis aleatórias $X_1,\ldots,X_n$, independentes e identicamente distribuídas, com a mesma distribuição de X. Com base nisso, faça o que se pede abaixo.

a) Vimos em sala de aula que uma boa forma de se estimar μ é através da *média amostral*, definida por

$$\overline{X}_n = \frac{X_1+\cdots+X_n}{n}.$$

Lembre-se que $\overline{X}_n$ também é uma variável aleatória, pois depende do sorteio de uma amostra da população. Portanto, calcule os valores de $\mathbb{E}[\overline{X}_n]$ e $\mathbb{V}(\overline{X}_n)$ e com base nisso justifique porque $\overline{X}_n$ pode ser considerado um "bom estimador" para μ. ***(1,5)***
Obs.: As aspas acima não contém nenhuma ironia, e somente indicam que não formalizamos completamente fatos sobre a qualidade de um estimador. Idem para as aspas do item b).

b) Agora, vamos tentar aferir algo mais quantitativo sobre a "qualidade" do estimador $\overline{X}_n$. Mostre que a probabilidade do estimador $\overline{X}_n$ se desviar do valor verdadeiro μ por mais do que c unidades é dada, aproximadamente, por $2\left[1-\Phi\left(\frac{c\sqrt{n}}{\sigma}\right)\right]$, ou seja, mostre que

$$\mathbb{P}\left(|\overline{X}_n-\mu|>c\right) \approx 2\left[1-\Phi\left(\frac{c\sqrt{n}}{\sigma}\right)\right],$$

onde Φ denota a função de probabilidade acumulada de uma distribuição normal padrão. ***(1,5)***

c) Deseja-se estimar o salário médio de uma população, na unidade monetária de salários mínimos. Pesquisas anteriores afirmam que o desvio padrão dos salários é dado por 1,1 salário mínimo. Qual deve ser o tamanho mínimo da amostra para garantir que tenhamos, com 99% de confiança, que o salário médio estimado não se afastará do salário médio "verdadeiro" por mais do que 1% de um salário mínimo. ***(1,0)***
Dica: Você já tem todo o cenário pronto no item b)! Adicionalmente, já fizemos exemplos bem semelhantes a esse em aula. Para encontrar informações sobre a acumulada da distribuição normal padrão recomendo utilizar o recurso do Geogebra que está no seguinte link: `https://www.geogebra.org/m/W9Nz53Ct`*. Por padrão, o aplicativo já abre com a densidade de uma normal padrão. Para visualizar a acumulada, clique no botão ao lado de onde está escrito "Normal", na parte inferior esquerda (ao posicionar o mouse em cima desse botão, irá aparecer escrito "Cumulative"). Para encontrar o valor que atinge determinada probabilidade de acordo com a acumulada da normal padrão, você pode colocar numericamente o valor na parte mais inferior do aplicativo ou deslizar o slider triangular no gráfico até encontrar o valor desejado.*

Questão 4: (QUESTÃO BÔNUS) Considere o espaço amostral $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$, e denote $\mathbb{P}(\{k\})$ por p_k, para $k = 1, \ldots, 8$, sendo tais valores definidos como abaixo:

$$
\begin{aligned}
p_1 &= \frac{8\alpha}{9} \\
p_2 = p_3 = p_4 &= \frac{7 - 16\alpha}{27} \\
p_5 = p_6 = p_7 &= \frac{1 + 8\alpha}{27} \\
p_8 &= \frac{1}{9}.
\end{aligned}
$$

a) Encontre um intervalo válido para α, ou seja, um intervalo no qual tenhamos todos os $p_k \geq 0$ e que tenhamos também $\sum_k p_k = 1$. ***(1,0)***

b) Defina os eventos abaixo:

$$
\begin{aligned}
A_1 &= \{2, 5, 6, 8\} \\
A_2 &= \{3, 5, 6, 8\} \\
A_3 &= \{4, 6, 7, 8\}.
\end{aligned}
$$

Estude condições sobre α para termos tais eventos independentes. ***(1,0)***