

Data: 13 de junho de 2023

Duração: 2 horas

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

Número: \_\_\_\_\_

### Instruções:

- As questões devem ser respondidas em folha de teste da ESTSetúbal (não se aceitam questões escritas a lápis) e os cálculos devem ser entregues num script do R (ficheiro com extensão .R) que deve ser entregue via Moodle.
- Na folha de teste não pode responder a diferentes questões numa mesma folha de resposta. Coloque o seu nome e número em todas as folhas de teste e no enunciado.
- Não é permitido partilhar ficheiros durante a prova nem utilizar o email ou outras formas de comunicação.
- Só é permitido aceder ao Moodle no início do teste para obter os dados e no fim do teste para enviar o script (ficheiro com extensão .R).
- Durante o teste o seu computador deverá estar sempre no R ou RStudio, não é permitido aceder a outros programas.
- É permitida a consulta do caderno/apontamentos e dos slides das aulas em versão papel.
- O abandono da sala só poderá efetuar-se decorridos 60 minutos a partir do início da prova e implica a entrega da mesma.
- **Justifique convenientemente todas as respostas na folha de teste da ESTSetúbal** e recorra ao software R apenas para efetuar os cálculos necessários.

### Questões:

1. Num jornal saiu a seguinte notícia: "Com o teletrabalho a facilitar a economia global, multinacionais recrutam cada vez mais profissionais de Tecnologias de Informação (TI) em Portugal, com salários mais elevados. Um problema para as tecnológicas nacionais, que correm risco de perder os melhores talentos. (...) Para as empresas que trabalham sobretudo para o mercado doméstico, esta questão é ainda mais sensível, pois vendem os seus produtos a preços "nacionais", mas pagam os seus recursos humanos a preços internacionais."

Com o objetivo de comprovar estas informações foi realizado um estudo através de questionário tendo sido recolhida uma amostra de profissionais portugueses da área de Tecnologias de Informação (TI) que trabalham em Portugal. No ficheiro "profTI.txt" que se encontra no Moodle, são apresentados os dados recolhidos referentes aos seguintes campos do questionário:

- Id = identificação do profissional de TI
- Salario = salário bruto mensal atual
- Be = se recebe algum benefício extra salário (0 = não, 1 = sim)
- In = salário bruto mensal inicial (quando entrou para a empresa)
- Exp = Experiência profissional (J = Júnior, S = Sénior)
- Genero = género do profissional de TI (0 = masculino, 1 = feminino)

- [2.0] (a) Calcule uma estimativa pontual para a percentagem de profissionais de TI que recebem benefícios extra salário e construa um intervalo de confiança a 96% para essa percentagem.

- [2.0] (b) De acordo com o Banco de Portugal, o salário médio mensal de entrada no mercado de trabalho em 2020 era de 1050 euros para um jovem com licenciatura. Com base no salário inicial dos profissionais de TI Juniores e recorrendo a um intervalo de confiança a 99%, comente a informação do Banco de Portugal.
- [2.0] (c) Segundo o Instituto Nacional de Estatística, a média mensal do salário bruto em Portugal no 2.º trimestre de 2022 ficou em 1439 euros. Para um nível de significância de 1%, teste se o salário médio atual dos profissionais de TI é superior à média dos salários em Portugal.
- [2.0] (d) Para um nível de significância de 10%, pode afirmar que existem diferenças significativas entre o salário inicial e o salário atual dos profissionais de TI Seniores do género feminino?
- [2.0] (e) Tipicamente as empresas não fazem discriminação de género em relação à concessão de benefícios extra salário. Para um nível de significância de 1% e com base nos dados que dispõe, diga, justificando, se concorda com esta afirmação. Se considerar adequado, quantifique a associação existente.
2. Para estudar a proporção de avarias devido a má utilização de um certo equipamento, o gerente de uma agência de reparações selecionou, aleatoriamente, 200 processos de reparação, os quais revelaram que em 30 desses casos a avaria foi consequência direta de má utilização.

- [2.0] (a) Interprete e teste as seguintes hipóteses:

$$H_0 : p = 0.16 \quad \text{contra} \quad H_1 : p < 0.16$$

Para realizar o teste recorra à região crítica e considere um nível de significância de 2%.

- [2.0] (b) Os equipamentos reparados nessa agência têm incorporado uma bateria, cuja duração considera-se que segue uma distribuição normal com um desvio padrão de 0.4 horas. Pretende-se estimar a duração média dessas baterias. Qual o número mínimo de baterias que se deve selecionar para uma amostra, de modo a garantir que se obtém um intervalo de confiança a 90% para a duração média das baterias com uma amplitude que não ultrapasse as 0.2 horas?
- (c) Com o objetivo de comparar a duração, em horas, de dois tipos de bateria desses equipamentos selecionou-se aleatoriamente duas amostras e registou-se a duração das baterias:

bateria do tipo XX:	7.5,	7.6,	7.7,	7.8,	7.9,	8.0,	8.1,	8.3
bateria do tipo YY:	7.1,	7.3,	7.7,	8.0,	8.3,	8.6		

- [1.0] i. Habitualmente supõe-se que a duração das baterias segue uma distribuição normal. Com base nos dados que dispõe e para cada um dos tipos de bateria, diga a partir de que nível de significância esta suposição não é válida.
- [2.0] ii. Sabendo que  $\sigma_{xx}$  refere-se às baterias do tipo XX e  $\sigma_{yy}$  às baterias do tipo YY, obteve-se o seguinte intervalo de confiança para  $\frac{\sigma_{xx}}{\sigma_{yy}}$ : ]0.20346; 0.94058[. Qual o grau de confiança utilizado?
- [1.0] iii. Calcule a margem de erro do intervalo de confiança da alínea anterior e dê sugestões para diminuir essa margem de erro.
- [2.0] iv. Teste, para um nível de significância de 10%, se, em média, é possível dizer que a duração das baterias do tipo XX ultrapassam a duração das baterias do tipo YY em mais de 30 minutos.

Fim do Teste