Lista nº 4 – Comparações de Médias - Respostas

1) Dados fornecidos pelo arquivo: *runtime.txt*.
   a) Teste das hipóteses H_0: $\mu_{Gloo} - \mu_{Cold} = 0$ *versus* H_1: $\mu_{Gloo} - \mu_{Cold} \neq 0$. Usando teste de Welch para duas amostras, tem-se $t_0 = 2{,}73$, com p-valor = 0,012. Há diferença forte de diferença entre as duas escolas. Um intervalo de confiança de 95% para as duas médias é [0,26; 1,91], ou seja, uma diferença de cerca de 0,3 a 1,9 segundos. Os métodos da teoria normal parecem se aplicáveis.
   b) Usando medianas, H_0: $\tilde{\mu}_{Gloo} - \tilde{\mu}_{Cold} = 0$ *versus* H_1: $\tilde{\mu}_{Gloo} - \tilde{\mu}_{Cold} \neq 0$. O teste de Mann-Whitney (Wilcoxon) dá um p-valor de 0,036, o qual fornece alguma evidência de uma diferença de escolas. Um intervalo de confiança aproximado de 95% para a diferença de medianas é [0,16; 1,92]
   c) Aqui tem-se um estudo observacional e não um experimento, logo não podemos provar causalidade, a saber, que os técnicos fazem diferença. Por exemplo, os melhores corredores poderiam ir para Glooscap. (Como você provaria que o treinamento faz diferença?)

2) Dados do arquivo *alfa.txt*.
   a) Usado o método de comparações emparelhadas, pois são medidas as mesmas marcas. Seja dif = alto – baixo. Então H_0: $\mu_{dif} = 0$ *versus* H_1: $\mu_{dif} > 0$. Usando um teste t de uma amostra, tem-se $t_0 = 2{,}01$, com p-valor (unilateral) = 0,37. Há evidência moderada apoiando a afirmação. A suposição de normalidade parece razoável.
   b) Não, pois as duas marcas têm diferenças negativas. O teste olha a diferença das médias, não diferenças individuais.
   c) Um gráfico de dispersão mostra que há uma relação fraca que parece ser quase inexistente para as sete observações mais próximas da origem.
   d) Através da aleatorização podemos tentar eliminar qualquer vício sistemático devido à ordem na qual os anúncios são vistos

3) Dados do arquivo *house.txt*.
   a) Estatísticas amostrais: $\bar{x}_{87} = 101{,}59$, $s_{87} = 36{,}11$, $\bar{x}_{89} = 134{,}37$, $s_{89} = 76{,}89$, $\bar{x}_{91} = 139{,}33$, $s_{91} = 66{,}19$. Há um aumento substancial na média amostral de novembro de 87 a Setembro de 89 e Agosto de 91. Há também um aumento substancial na dispersão. A partir dos gráficos de pontos vemos que isto é possivelmente devido a umas poucas moradias mais caras em 1989 e 1991. O box-plot mostra tendências semelhantes, embora as diferenças não aparecem ser tão óbvias devido à escala vertical comprimida.
   b) O teste F dá $f_0 = 3{,}65$, com um p-valor de 0,030, fornecendo evidência moderada contrária à igualdade das três médias. Os intervalos de confiança individuais de 95% indicam que 87 é muito diferente de 89 e 91, que são semelhantes.
   c) Usando comparações de Fisher para cada par, temos 91 – 87: [5,1; 70,3]; 89 – 87: [3,5; 62,1]; 91 – 89: [-29,0; 38,9]
   d) Verifica-se que $s_{89} > 2s_{87}$. Há valores atípicos presentes e os dados de 89 são claramente assimétricos. A teoria normal é, portanto, suspeita.
   e) Exceto por possíveis valores atípicos, os gráficos de pontos indicam que a maior parte da assimetria parece ter sido removida e as dispersões são muito semelhantes.
   f) Usando de um teste F com os logaritmos dos dados, obtemos $f_0 = 4{,}00$, com um p-valor de 0,022, logo nossa conclusão não se altera. Os dados são ainda

assimétricos, como visto a partir do histograma dos resíduos e a leve curvatura no seu gráfico de probabilidade normal. O teste W de normalidade tem um p-valor < 0,01. Contudo os desvios padrão são agora semelhantes. Novamente concluímos que há um aumento significante nos preços de moradia de 1987 a 1989 e pouca mudança de 1989 a 1991

g) Um aumento de preço médio de moradia não implica que todos os preços de moradia cresçam; da mesma forma alguns decrescem. Além disso, qualquer aumento na média pode ser devido apenas a serem vendidas algumas poucas moradias caras. Esses comentários se aplicariam a todas as moradias. Precisaríamos olhar as moradias vendidas mais de uma vez ou, se houver poucas nessa categoria, compare moradias de avaliações semelhantes.

4)

a) (i) Uma amostra. Intervalo de confiança (pois não foi dito o que significa “efetivo”). (ii) as porcentagens são aproximadamente normais com desvios padrão iguais. (iii) nenhum placebo é usado para comparação. Também, alguns pacientes terão mais dor de cabeça do que outros de modo que as porcentagens (binomial) terão desvios padrão diferentes.

b) (i) Duas amostras independentes. Intervalo de confiança. (ii) o conjunto de dados para cada método é normalmente distribuído e os dois conjuntos são independentes. (Também precisamos utilizar um desenho completamente aleatorizado.) (iii) Variabilidade na fertilidade, por exemplo, dos canteiros, o que pode vir a ser confundido com a diferença dos métodos.

c) (i) mais de duas amostras independentes. Intervalos de confiança. (ii) Suponha que o número de capturados de cada cor é normalmente distribuído e que os quatros desvios padrão são todos iguais. Suponha também que as quatro amostras são independentes. (Precisamos de ter algum método aleatorizado, tal como um desenho de bloco aleatorizado, para alocar cada placa.) (iii) Pode haver uma variação no número de besouros nas diferentes partes do campo.