UFV- CCE - DET

EST 105 - 3^a avaliação - 1^0 semestre de 2019 - 29/junho/19

Nome: ______________________________ **Matrícula:**________________

Assinatura: ________________________. Favor apresentar documento com foto.

- São 5 questões, tabelas e formulário em páginas numeradas de 1 a 9, total de 40 pontos, FAVOR CONFERIR ANTES DE INICIAR.

- ATENÇÃO: Assinale (X) em qual turma está matriculado (sua nota será divulgada no sistema SAPIENS).

```
-------------------------------------------------------------------------
      TURMA    HORÁRIO         SALA   PROFESSOR
-------------------------------------------------------------------------
(   )   T1 3ª 8-10   5ª 10-12  PVB300 Cecon
(   )   T2 3ª 10-12 6ª 8-10    PVB104 Carol
(   )   T3 3ª 14-16 5ª 16-18   PVB100 Camila/Carol
(   )   T4 2ª 14-16 4ª 16-18   PVB107 Moysés
(   )   T5 4ª 18:30-20:10 6ª 20:30-22:10  PVB204 Eduardo
(   )   T6 4ª 14-16 6ª 16-18   PVA361  CHOS - coordenador
(   )   T7 2ª 16-18 5ª 14-16   PVB307 Camila
(   )   T8 2ª 20:30-22:10 5ª 18:30-20:10 PVA361 Eduardo
-------------------------------------------------------------------------
```

- Interpretar corretamente as questões é parte da avaliação, portanto não é permitido questionamentos durante a prova !

- É OBRIGATÓRIO APRESENTAR OS CÁLCULOS organizadamente, para ter direito à revisão.

- PODE UTILIZAR A CALCULADORA, porém mostre os valores utilizados na fórmula.

- BOA SORTE e BOA PROVA !!!.

FORMULÁRIO

$$f(x|y) = \frac{f(x,y)}{h(y)}, \quad h(y) = \int f(x,y)\ dx, \qquad f(y|x) = \frac{f(x,y)}{g(x)}, \quad g(x) = \int f(x,y)\ dy$$

$$P(x|y) = \frac{P(x,y)}{P(y)}, \quad P(y) = \sum_x P(x,y), \qquad P(y|x) = \frac{P(x,y)}{P(x)}, \quad P(x) = \sum_y P(x,y)$$

$$\text{Para} \;\; k = 1, 2, \ldots, n < \infty \qquad E(X^k) = \sum_x x^k P(x) \quad \text{ou} \quad E(X^k) = \int x^k f(x) dx$$

$$E(XY) = \sum_x \sum_y xyP(x,y) \quad \text{ou} \quad E(XY) = \int \int xyf(x,y) dxdy$$

$$\text{cov}(X,Y) = E(XY) - E(X)E(Y), \quad \rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sqrt{V(X)V(Y)}}, \quad V(X) = E(X^2) - [E(X)]^2$$

$$X \ \sim \ N(\mu, \sigma^2), \quad E(X) = \mu \ \text{ e } \ V(X) = \sigma^2 \quad Z = \frac{X-\mu}{\sigma}, \quad Z \ \sim \ N(0,\ 1)$$

$$P(x) = \binom{N}{x} p^x (1-p)^{N-x} \qquad \binom{N}{x} = \frac{N!}{x!(N-x)!} \qquad E(X) = Np \;\; V(X) = Np(1-p)$$

$$P(x) = \frac{e^{-m} m^x}{x!} \qquad E(X) = V(X) = m$$

$$Z = \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

$$t = \frac{\left(\overline{X}_A - \overline{X}_B\right)}{\sqrt{S^2\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} \qquad S^2 = \frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}$$

Tabela 1: Áreas de uma distribuição normal padrão entre $z = 0$ e um valor positivo de z. As áreas para os valores de z negativos são obtidas por simetria.

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4006	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

Adaptada de Costa Neto, P. L. O. Estatística, Editora Edgard Blucher.

Tabela 2: Valores positivos t na distribuição t_n de Student com n graus de liberdade em níveis de 10% a 0,1% de probabilidade $= 2 \times P\,(t_n \geq t)$, tabela bilateral.

	nível de probabilidade bilateral					
n	10%	5%	2%	1%	0,5%	0,1%
1	6,31	12,71	31,82	63,66	127,32	636,62
2	2,92	4,30	6,97	9,92	14,09	31,60
3	2,35	3,18	4,54	5,84	7,45	12,94
4	2,13	2,78	3,75	4,60	5,60	8,61
5	2,02	2,57	3,37	4,03	4,77	6,86
6	1,94	2,45	3,14	3,71	4,32	5,96
7	1,90	2,36	3,10	3,50	4,03	5,41
8	1,86	2,31	2,90	3,36	3,83	5,04
9	1,83	2,26	2,82	3,25	3,69	4,78
10	1,81	2,23	2,76	3,17	3,58	4,59
11	1,80	2,20	2,72	3,11	3,50	4,44
12	1,78	2,18	2,68	3,06	3,43	4,32
13	1,77	2,16	2,65	3,01	3,37	4,22
14	1,76	2,14	2,62	2,98	3,33	4,14
15	1,75	2,13	2,60	2,95	3,29	4,07
16	1,75	2,12	2,58	2,92	3,25	4,02
17	1,74	2,11	2,57	2,90	3,22	3,97
18	1,73	2,10	2,55	2,88	3,20	3,92
19	1,73	2,09	2,54	2,86	3,17	3,88
20	1,73	2,09	2,53	2,84	3,15	3,85
21	1,72	2,08	2,52	2,83	3,14	3,82
22	1,72	2,07	2,51	2,82	3,12	3,79
23	1,71	2,07	2,50	2,81	3,10	3,77
24	1,71	2,06	2,49	2,80	3,09	3,75
25	1,71	2,06	2,49	2,79	3,08	3,73
26	1,71	2,06	2,48	2,78	3,07	3,71
27	1,70	2,05	2,47	2,77	3,06	3,69
28	1,70	2,05	2,47	2,76	3,05	3,67
29	1,70	2,04	2,46	2,76	3,04	3,66
30	1,70	2,04	2,46	2,75	3,03	3,65
40	1,68	2,02	2,42	2,70	2,97	3,55
60	1,67	2,00	2,39	2,66	2,92	3,46
120	1,65	1,98	2,36	2,62	2,86	3,37
∞	1,65	1,96	2,33	2,58	2,81	3,29

Adaptada de Frederico Pimentel Gomes, Curso de Estatística Experimental, 12^a ed.

1.(8 pontos) Seja (X, Y) uma variável aleatória bidimensional com a seguinte distribuição conjunta de probabilidades,

X	Y			$P(y)$
	1	3	5	
0	0,10	0,10	0,05	0,25
1	0,20	0,20	0,10	0,50
2	0,10	0,10	0,05	0,25
$P(y)$	0,40	0,40	0,20	1

Seja $W = 3X - 2Y + 5$, aplique as propriedades e calcule:

a.(3 pts) $E(W)$, o valor médio de W.

$$\begin{aligned} E(W) &= E(3X - 2Y + 5) = E(3X) - E(2Y) + E(5) \\ &= 3E(X) - 2E(Y) + 5 \end{aligned}$$

como

$$\begin{aligned} E(X) &= \sum_x xP(x) = 0 \cdot 0,25 + 1 \cdot 0,50 + 2 \cdot 0,25 = 1 \\ E(Y) &= \sum_y yP(y) = 1 \cdot 0,40 + 3 \cdot 0,40 + 5 \cdot 0,20 = 2,6 \end{aligned}$$

$$E(W) = 3 \cdot 1 - 2 \cdot 2,6 + 5 = 2,8$$

b.(5 pts) $V(W)$, a variância de W.

$$\begin{aligned} V(W) &= V(3X - 2Y + 5) = V(3X - 2Y) \\ &= V(3X) + V(2Y) - 2Cov(3X, 2Y) \\ &= 9V(X) + 4V(Y) - 2 \times 3 \times 2Cov(X, Y) \\ &= 9V(X) + 4V(Y) - 12Cov(X, Y) \end{aligned}$$

como $P(x, y) = P(x)P(y)$, $\forall x$ e y, temos que X e Y são independentes e, desta forma $Cov(X, Y) = 0$, além disso

$$\begin{aligned} E\left(X^2\right) &= \sum_x x^2 P(x) = 0^2 \cdot 0,25 + 1^2 \cdot 0,50 + 2^2 \cdot 0,25 = 1,5 \\ E\left(Y^2\right) &= \sum_y y^2 P(y) = 1^2 \cdot 0,40 + 3^2 \cdot 0,40 + 5^2 \cdot 0,20 = 9 \end{aligned}$$

assim

$$\begin{aligned} V(X) &= E\left(X^2\right) - (E(X))^2 = 1,5 - (1)^2 = 0,5 \\ V(Y) &= E\left(Y^2\right) - (E(Y))^2 = 9 - (2,6)^2 = 2,24 \end{aligned}$$

finalmente

$$V(W) = 9 \cdot 0,5 + 4 \cdot 2,24 = 13,46$$

2.(8 pontos) Seja $f(x,y)$ uma função densidade de probabilidade conjunta dada por,

$$f(x,y)=\begin{cases} 4xy & , \quad 0\le x\le 1 \ \text{ e } \ 0\le y\le 1 \\ \\ 0 & , \quad \text{para outros valores } \ x,y \end{cases}$$

Verifique se X e Y são variáveis aleatórias contínuas independentes. Responda SIM ou NÃO e apresente os cálculos que justificam sua resposta.

$$g\left(x\right)=\int f\left(x,y\right)dy=\int_0^1 4xydy=4x\left.\frac{y^2}{2}\right|_{y=0}^{1}=2x,$$

$$h\left(y\right)=\int f\left(x,y\right)dx=\int_0^1 4xydy=4y\left.\frac{x^2}{2}\right|_{x=0}^{1}=2y.$$

Como $f\left(x,y\right)=g\left(x\right)h\left(y\right)$ então X e Y são variáveis aleatórias independentes.

3.(8 pontos) Assuma que 2% dos fregueses desistem de fazer o pagamento das compras quando estão no caixa e são informados do valor total a pagar.Seja X a variável aleatória que represente o número de fregueses que se comportam de tal modo em uma amostra aleatória de 300 fregueses. Pede-se: Calcule $P(X \geq 3)$, a probabilidade de pelo menos 3 fregueses desistirem,

a.(4 pts) considerando-se o modelo Binomial.

$N = 300$, $p = 0,02$ e $P(X = x) = \binom{N}{x} p^x (1-p)^{N-x}$. Temos pelo complemento que

$$\begin{aligned}
P(X \geq 3) &= 1 - P(X < 3) = 1 - (P(X = 0) + P(X = 1) + P(X = 2)) \\
&= 1 - \left[\binom{300}{0} \times 0,02^0 \times 0,98^{300} + \binom{300}{1} \times 0,02^1 \times 0,98^{299} \right. \\
&+ \left. \binom{300}{2} \times 0,02^2 \times 0,98^{298} \right] \\
&= 1 - [0,0023 + 0,0143 + 0,0436] \\
&= 1 - 0,0602 \\
&= 0,9398 \quad (\approx 94\%)
\end{aligned}$$

b.(4 pts) Considerando-se o modelo Poisson.

$m = Np = 300 \times 0,02 = 6$ e $P(X = x) = \frac{e^{-m} m^x}{x!}$. Temos pelo complemento que

$$\begin{aligned}
P(X \geq 3) &= 1 - P(X < 3) = 1 - (P(X = 0) + P(X = 1) + P(X = 2)) \\
&= 1 - \left(\frac{e^{-6} 6^0}{0!} + \frac{e^{-6} 6^1}{1!} + + \frac{e^{-6} 6^2}{2!} \right) = 1 - \left(e^{-6} + 6e^{-6} + \frac{36e^{-6}}{2} \right) \\
&= 1 - 25e^{-6} = 1 - 0,062 = 0,938 \quad (\approx 94\%)
\end{aligned}$$

De outra forma:

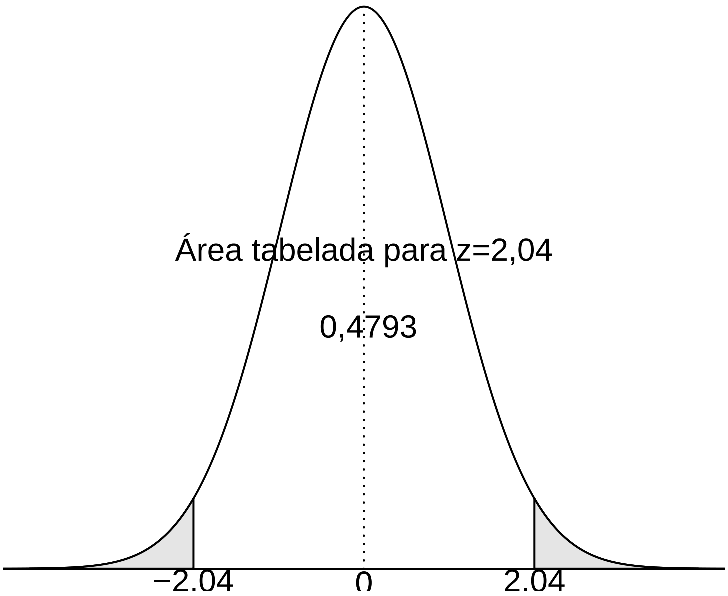
$$\begin{aligned}
P(X \geq 2) &= 1 - e^{-m} \left(\frac{m^0}{0!} + \frac{m^1}{1!} + + \frac{m^2}{2!} \right) \\
&= 1 - e^{-6} (1 + 6 + 18) = 1 - 25e^{-6} \\
&= 1 - 0,062 = 0,938 \quad (\approx 94\%)
\end{aligned}$$

4.(8 pontos) Uma máquina está regulada para fornecer $\mu = 500$g por pacote e seja $\sigma^2 = 25\text{g}^2$. Uma amostra aleatória de tamanho $n = 36$ pacotes forneceu peso médio $\overline{X} = 501,7$g, teste $H_0 : \mu = 500$ (a máquina está regulada) contra $H_1 : \mu \neq 500$ (a máquina está desregulada) conforme os itens a seguir.

a.(3 pts) Valor calculado.

$$z_{\text{cal}} = z_0 = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{501,7 - 500}{\sqrt{\frac{25}{36}}} = \frac{1,7}{\frac{5}{6}} = 2,04$$

b.(2 pts) Valor-p, faça um desenho ilustrativo.



Valor-p$= 2 \cdot (0,5 - 0,4793) = 0,0414 \quad (4,14\%)$.

c.(3 pts) Adote um nível de significância à sua escolha e conclua.

Se $\alpha \geq 4,14\%$ rejeita-se H_0, ou seja, H_0 é rejeitada para, por exemplo, $\alpha = 5\%$ e não é rejeitada para $\alpha = 1\%$.

$\alpha = 1\%$, $z_{\text{tabelado}} = 2,57$ ou 2,58

$\alpha = 5\%$, $z_{\text{tabelado}} = 1,96$

Regra geral: se valor-p$\leq \alpha$ então rejeita-se H_0.

5.(8 pontos) Acredita-se que um programa escrito em C++ deva rodar, em média, mais rapidamente do que o mesmo programa escrito em FORTRAN ($\mu_C < \mu_F$). Na tabela a seguir são apresentados os tempos (minutos) obtidos quando os programas foram aplicados em amostras representativas de tarefas bastante complexas. Adote 5% como nível de significância e responda aos itens a seguir. Note que o tempo da tarefa 5 em C++ não foi registrado.

	Tarefas					
Programa	1	2	3	4	5	6
C++	0,8	4,4	1,3	3,9	–	5,0
FORTRAN	3,3	7,5	1,3	6,3	8,1	5,2

a.(1 pt) Hipóteses estatísticas:

$$\begin{cases} H_0: & \mu_C = \mu_F \\ H_1: & \mu_C < \mu_F. \end{cases}$$

b.(2 pts) Valor tabelado:

$t_{tab} = t_{(9,5\%)} = 1,83$, deve-se olhar 10% na tabela bilateral.

c.(2 pts) Valor calculado:

$$S_c^2 = \frac{(5-1)\,3,617 + (6-1)\,6,7377}{5+6-2} = \frac{48,1565}{9} = 5,3507.$$

$$t_{\text{cal}} = \frac{3,08 - 5,2833}{\sqrt{5,3507\left(\frac{1}{5}+\frac{1}{6}\right)}} = -1,57.$$

d.(1 pt) Decisão do teste:

Como $|-1,57| < 1,83$, não se rejeita H_0 a 5% de significância.

e.(2 pts) Interpretação prática do resultado do teste para o problema:

Não há evidências (ou indícios) de que o programa em C++ rode mais rapidamente (em média) do que em FORTRAN.