PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TRANSPORTES E GESTÃO TERRITORIAL – PPGTG
DEPARTAMENTO DE ENGENHARIA CIVIL – ECV

DISCIPLINA: TGT410026 – FUNDAMENTOS DE ESTATÍSTICA

5ª AULA: DISTRIBUIÇÃO DE PROBABILIDADE

Nesta aula, continuamos o estudo da probabilidade, introduzindo o conceito de variável aleatória e de distribuição de probabilidade.

Nas aulas anteriores discutimos o conceito de experimento e seus resultados experimentais concomitantes. Uma ***variável aleatória*** fornece um meio para se descrever resultados experimentais usando-se valores numéricos. Em suma, ***uma variável aleatória é uma descrição numérica do resultado de um experimento***.

Uma variável aleatória associa um valor numérico a cada resultado experimental possível. O valor numérico da variável aleatória depende, portanto, do resultado do experimento. A variável aleatória é classificada em discreta ou contínua, dependendo dos valores numéricos que ela assume.

VARIÁVEIS ALEATÓRIAS DISCRETAS – uma variável que pode assumir tanto um valor finito de valores como uma sequência infinita de valores – tais como 0, 1, 2, ... – denomina-se ***variável aleatória discreta (VAD)***.

Exemplos: seja o caso de um concurso para perito-contador. O concurso é composto de quatro etapas. Pode-se definir uma variável aleatória como x = *número de etapas em que ele foi aprovado*. Trata-se de uma VAD porque ela pode assumir os valores 0, 1, 2, 3 ou 4.

Considere agora o experimento de observar o número de carros que chegam a um posto de pedágio. A variável de interesse seria x = *número de carros que chegam ao pedágio em um dia*. Trata-se de uma VAD, pois ela pode assumir qualquer valor da sequência de números inteiros 0, 1, 2, ... e assim por diante. Portanto, x é uma VAD que assume um dos valores dessa sequência infinita.

Muitos experimentos apresentam resultados que são naturalmente descritos por valores numéricos, mas alguns não o são. Como exemplo, seja o caso de uma pesquisa onde uma das questões seja solicitar a um indivíduo que relembre a mensagem de um recente comercial de TV. Obviamente as duas respostas possíveis seriam: o indivíduo não é capaz de lembrar-se da mensagem e o indivíduo é capaz de recordar-se da mensagem. Para descrever esses resultados de forma numérica, vamos definir uma VAD da seguinte maneira: $\boldsymbol{x = 0}$ se o indivíduo não consegue lembrar-se da mensagem e $\boldsymbol{x = 1}$, se o indivíduo consegue relembrar a mensagem.

Outros exemplos de VAD:

Experimento	**Variável Aleatória (x)**	**Valores possíveis para a VA**
Contatar 5 clientes	Nº de clientes que fazem um pedido de compra	0, 1, 2, 3, 4, 5
Inspecionar o embarque de 50 TVs	Nº de TVs defeituosas	0, 1, 2, 3, ..., 49, 50
Operar um restaurante num dia	Nº de clientes	0, 1, 2, 3, ...
Vender um automóvel	Gênero do cliente	0 se for masculino; 1 se for feminino

VARIÁVEIS ALEATÓRIAS CONTÍNUAS – é a variável que pode assumir qualquer valor numérico em um intervalo ou em um conjunto de intervalos. Resultados experimentais que se baseiam em escalas de medidas como o tempo, peso, distância, temperatura, etc... são descritos por meio de variáveis aleatórias contínuas – VAC.

Como exemplo, seja o experimento de monitoração das chamadas telefônicas feitas ao escritório de reclamação de uma importante companhia de seguros. Suponha que a variável de interesse seja *x = o tempo em minutos entre as chamadas consecutivas*. Essa variável aleatória pode assumir qualquer valor no intervalo $\boldsymbol{x \geq 0}$. Realmente, um número infinito de valores é possível para *x*, incluindo valores tais como: 1,46 min, 4,3333 min, 2,731 min e assim por diante.

Outro exemplo, considere um trecho da rodovia BR-101 de 60km ao norte de Florianópolis. Para o serviço de emergência de ambulâncias localizado em Florianópolis, podemos definir a variável aleatória *x = nº de quilômetros até o local do próximo acidente de trânsito ao longo desse trecho da BR-101*. Nesse caso, *x* seria uma variável aleatória contínua que assume qualquer valor no intervalo $\boldsymbol{0 \leq x \leq 60}$.

Outros exemplos de VAC:

Experimento	**Variável Aleatória (x)**	**Valores possíveis para a VA**
Operar um banco	Tempo em minutos entre as chegadas dos clientes	$X \geq 0$
Encher uma lata de refrigerante de 343 ml	Quantidade em ml	$0 \leq x \leq 343$
Construir uma nova biblioteca	Porcentagem de conclusão do projeto depois de 6 meses	$0 \leq x \leq 100$
Testar um novo processo químico	Temperatura quando ocorre a reação desejada – mín de 65º e máx de 100ºC	65ºC ≤ x ≤ 100ºC

DISTRIBUIÇÃO DE PROBABILIDADE

Distribuição de probabilidade de uma variável aleatória é um modelo matemático que descreve como as probabilidades estão distribuídas sobre os valores da variável aleatória. Essa distribuição de probabilidades é definida por uma função denominada ***função probabilidade***, denotada por *f(x)*.

Para uma variável aleatória discreta *x* a *f(x)* fornece a probabilidade correspondente a cada um dos valores da variável aleatória.

No desenvolvimento de uma função probabilidade para qualquer VAD, as duas condições seguintes devem ser satisfeitas:

$$f(x) \geq 0$$
$$\sum f(x) = 1$$

Para fixar, vamos usar como ilustração o comportamento de uma VAD e sua distribuição de probabilidades nas vendas ocorridas na DiCarlo Automóveis nos últimos 300 dias de operação. Os dados de vendas mostram: 54 dias sem vendas de automóveis; 117 dias com um automóvel vendido; 72 dias com dois automóveis vendidos; 42 dias com três automóveis vendidos; 12 dias com quatro automóveis vendidos e 3 dias com cinco automóveis vendidos. Seja o experimento de selecionar um dia de operação da DiCarlo Automóveis. Definimos a variável de interesse como *x = nº de automóveis vendidos durante um dia*.

Sabemos que *x* pode assumir os valores 0, 1, 2, 3, 4 ou 5. Na notação da função probabilidade, *f(0)* fornece a probabilidade de 0 automóveis vendidos, *f(1)* fornece a probabilidade de 1 automóvel vendido, e assim por diante.

x	***f(x)***
0	*f(0) = 54/300 = 0,18*
1	*f(1) = 117/300 = 0,39*
2	*f(2) = 72/300 = 0,24*
3	*f(3) = 42/300 =0,14*
4	*f(4) = 12/300 =0,04*
5	*f(5) = 3/300 =0,01*
	Total: 1,00

Vemos na tabela que todos os valores de *f(x)* satisfazem a primeira condição, ou seja, os valores de f(x) são maiores que zero; ademais a soma de todas as probabilidade é igual a 1 o que atende a segunda condição da função probabilidade; assim, a função probabilidade da DiCarlo Automóveis é uma função probabilidade discreta válida.

Aplicações:

1) Uma variável aleatória *x* possui a seguinte distribuição de probabilidade:

x	***f(x)***
20	0,25
25	0,15
30	0,25
35	0,40

a) essa distribuição de probabilidade é válida? Explique.
b) qual é a probabilidade de *x* ser igual a 30?
c) qual é a probabilidade de *x* ser menor ou igual a 25?
d) qual é a probabilidade de *x* ser maior que 30?

2) Os dados a seguir foram coletados contando-se o número de salas de cirurgia em uso num Hospital em um período de 25 dias: em três dos dias somente uma sala de cirurgia foi usada; em cinco dos dias duas salas foram usadas; em oito dos dias três salas foram usadas e em quatro dos dias todas as quatro salas de cirurgia do Hospital foram utilizadas.
a) use a abordagem da frequência relativa para construir a distribuição de probabilidade correspondente ao número de salas de cirurgia em uso em qualquer dia do período.
b) desenhe um gráfico da distribuição de probabilidade (Dica: *f(x)* no eixo vertical; *x* no eixo horizontal).
c) mostre que sua distribuição de probabilidade satisfaz as condições necessárias a uma distribuição de probabilidade de variável discreta válida.

3) Um psicólogo determinou que o número de sessões necessárias para conquistar a confiança de um novo paciente pode ser 1, 2 ou 3. Seja x uma variável aleatória que indica o nº de sessões necessárias para conquistar a confiança do paciente. A seguinte função de probabilidade foi proposta:

$f(x) = x / 6$ para x = 1, 2 ou 3.

a) essa é uma função probabilidade válida? Explique.

b) qual é a probabilidade de serem necessárias exatamente duas sessões para conquistar a confiança do paciente?

c) qual é a probabilidade de serem necessárias *pelo menos* duas sessões para conquistar a confiança do paciente?

4) Os dados a seguir são de uma distribuição de probabilidade parcial referente ao lucro projetado de uma empresa – *x = lucro em milhares de dólares* – para o primeiro ano de operação (o valor negativo denota um prejuízo).

x	*f(x)*
-100	0,10
0	0,20
50	0,30
100	0,25
150	0,10
200	?

a) qual é o valor adequado para *f(200)*? Dê sua interpretação do valor.

b) qual é a probabilidade de a empresa ser rentável?

c) qual a probabilidade de a empresa alcançar *pelo menos* US$ 100 mil?

5) O diretor de um Colégio avaliou subjetivamente uma distribuição de probabilidade para *x*, equivalente ao número de matriculandos, da seguinte maneira:

x	*f(x)*
1.000	0,15
1.100	0,20
1.200	0,30
1.300	0,25
1.400	0,10

a) a distribuição é válida? Explique.

b) qual a probabilidade de 1.200 estudantes ou menos se matricularem?

VALOR ESPERADO DE UMA VARIÁVEL DISCRETA E SUA DISPERSÃO

Em muitas aplicações pode surgir o interesse em, com base em algum conhecimento, efetuar uma estimativa futura, mesmo que seja uma estimativa preliminar. Isso é possível com o cálculo do valor esperado da variável aleatória em estudo.

O ***valor esperado (E(x))***, ou ***média (μ)***, de uma variável aleatória é a medida da posição central da variável aleatória. Ou seja, é a média dos valores da distribuição da VA ao longo de sua distribuição.
Analiticamente temos:

$$E(x) = \mu = \sum x.f(x)$$

Não obstante o valor esperado fornecer o valor médio da VA, frequentemente necessitamos de uma medida de variabilidade ou de dispersão. Tal como usamos a variância para sintetizar a variabilidade no conjunto de dados, usamos agora a variância para sintetizar a variabilidade nos valores da variável aleatória.
Analiticamente temos:

$$Var(x) = \sigma^2 = \sum (x-\mu)^2.f(x)$$

A raiz quadrada positiva da variância é o desvio padrão $\sigma = \sqrt{\sigma^2}$, frequentemente usado como medida de dispersão, pois possui unidade de medida idêntica a da VA.

Aplicação:
1) A tabela a seguir mostra o cálculo do valor esperado para o número de automóveis vendidos durante um dia pela DiCarlo Automóveis.

x	*f(x)*	*x.f(x)*
0	0,18	0.(0,18) = 0,00
1	0,39	1.(0,39) = 0,39
2	0,24	2.(0,24) = 0,48
3	0,14	3.(0,14) = 0,42
4	0,04	4.(0,04) = 0,16
5	0,01	5.(0,01) = 0,05
		$E(x) = \mu = \sum x.f(x) = 1,50$

Valor esperado é de 1,5 automóveis por dia. Isso significa que, embora seja possível a realização de 0, 1, 2, 3, 4 ou 5 vendas de automóveis em qualquer um dos dias, ao longo do tempo a DiCarlo Automóveis pode prever a venda de uma média de 1,50 automóvel por dia. Assim, num período de 30 dias de operação da DiCarlo Automóveis, pode-se usar o valor esperado para prever vendas mensais médias de 30.(1,5) = 45 automóveis.

E a incerteza na previsão? Para isso calculamos o desvio padrão dos valores da VA. A tabela a seguir mostra o procedimento de cálculo tomando os dados da DiCarlo Automóveis.

x	$x - \mu$	$(x - \mu)^2$	$f(x)$	$(x - \mu)^2.f(x)$
0	-1,5	2,25	0,18	2,25.0,18 = 04050
1	-0,50	0,25	0,39	0,25.0,39 = 0,0975
2	0,50	0,25	0,24	0,25.0,24 = 0,0600
3	1,50	2,25	0,14	2,25.0,14 = 0,3150
4	2,50	6,25	0,04	6,25.0,04 = 0,2500
5	3,50	12,25	0,01	12,25.0,01 = 0,1225
				$\sum(x - \mu)^2.f(x) = 1,2500$

A variância é igual a 1,25 e o desvio padrão é igual a $\sqrt{1,25}$ = 1,0.
Assim, a incerteza na previsão de vendas num dia qualquer é:
$\mu \pm \sigma$ = 1,5 ± 1,0 = [0,5 a 2,5]; adotando-se o valor inferior do intervalo, num período de 30 dias de operação da DiCarlo Automóveis poder-se-ia esperar a venda de apenas 0,5.30=15 automóveis; por outro lado, adotando-se o limite superior do intervalo, no mesmo período de operação da DiCarlo Automóveis poder-se-ia esperar uma venda de 2,5.30=75 automóveis.
Essa análise é importante para o gerente da DiCarlo Automóveis, pois auxiliaria a programar o estoque mensal, por exemplo.

DISTRIBUIÇÃO DE PROBABILIDADES USUAIS

As distribuições mais utilizadas em estudos estatísticos são: a binomial, de Poisson e a hipergeométrica para as variáveis discretas. Para as variáveis contínuas a distribuição mais utilizada é a distribuição normal ou de Gauss.

DISTRIBUIÇÃO BINOMIAL

A distribuição de probabilidade binomial é uma distribuição de probabilidade discreta que está associada a um experimento de múltiplas etapas, denominado *experimento binomial*.

Um experimento é dito binomial se atende as quatro propriedades a seguir:

- O experimento consiste em uma sequência de ***n*** ensaios ***idênticos***;

- Dois resultados são possíveis em cada ensaio: referimo-nos a um como um ***sucesso*** e ao outro como ***fracasso***;
- A probabilidade de um sucesso, denotado por ***p***, não se modifica de ensaio para ensaio; consequentemente a probabilidade de um fracasso, denotado por (***1 – p***) também se altera de ensaio para ensaio;
- Os ensaios são ***independentes***.

“Em um experimento binomial o ***interesse sempre***, é o número de sucessos que ocorrem nos ***n*** ensaios.”

FUNÇÃO PROBABILIDADE BINOMIAL – *f(x)*

A função probabilidade binomial *f(x)* é uma equação que fornece o *número de resultados de um experimento binomial com x sucessos* e também a probabilidade referente a cada sequência envolvendo *x sucessos*.

$$f(x) = C_{n,x} \,.\, p^{x}.(1-p)^{(n-x)}$$

onde $C_{n,x}$ é a combinação dos *n* ensaios com *x* sucessos; *p* é a probabilidade de sucesso em qualquer dos ensaios.

Quando o número de ensaios (*n*) é grande o cálculo das probabilidades é facilitado pelo uso da tabela de probabilidades binomiais. Para uso da tabela binomial devemos estar atentos aos parâmetros da distribuição binomial: ***n*** – nº de ensaios e ***p*** a probabilidade de sucesso.

Tabela: Valores selecionados da tabela de probabilidades binomiais

		p									
n	***x***	**0,05**	**0,10**	**0,15**	**0,20**	**0,25**	**0,30**	**0,35**	**0,40**	**0,45**	**0,50**
9	**0**	0,6302	0,3874	0,2316	0,1342	0,0751	0,0404	0,0207	0,0101	0,0046	0,0020
	1	0,2985	0,3874	0,3679	0,3020	0,2253	0,1556	0,1004	0,0605	0,0339	0,0176
	8					0,0001	0,0004	0,0013	0,0035	0,0083	0,0176
	9							0,0001	0,0003	0,0008	0,0020
10	**0**	0,5987	0,3487	0,1969	0,1074	0,0563	0,0282	0,0135	0,0060	0,0025	0,0010
	1	0,3151	0,3874	0,3474	0,2684	0,1877	0,1211	0,0725	0,0403	0,0207	0,0098
	2	0,0746	0,1937	0,2759	0,3020	0,2816	0,2335	0,1757	0,1209	0,0763	0,0439
	5	0,0001	0,0015	0,0085	0,0264	0,0584	0,1029	0,1536	0,2007	0,2340	0,2461
	7			0,0001	0,0008	0,0031	0,0090	0,0212	0,0425	0,0746	0,1172
	9						0,0001	0,0005	0,0016	0,0042	0,0098

Quando ***p > 50*** a mesma tabela pode ser utilizada; para isso basta calcular a probabilidade de ***n-x*** fracassos com ***p´= 1 - p***.

EXEMPLOS:

1) Considere um vendedor de seguros que visita dez famílias selecionadas aleatoriamente. O resultado associado a cada visita é classificado como um sucesso se a família comprar uma apólice de seguro, e como um fracasso se a família não comprar. Por experiência o vendedor estima que a probabilidade de uma família selecionada aleatoriamente comprar uma apólice de seguro é igual a 0,10. Verificar se o experimento se enquadra como um experimento binomial.

2) Considerar as decisões de compra dos próximos três clientes que entram na loja de roupas Renner. Com base em sua experiência o gerente da loja estima que a probabilidade de qualquer dos clientes comprar é de 0,30. Qual a probabilidade de dois dos próximos três clientes realizarem uma compra.

a) Verifique as quatro exigências para um experimento binomial.

b) Desenhe um diagrama em árvore e verifique os resultados experimentais possíveis (adote S para sucesso de venda; F para fracasso).

c) Determine o número de resultados experimentais que resultam em exatamente x sucessos em n ensaios (utilize a fórmula).

d) Calcule as probabilidades para 0, 1, 2 e 3 sucessos.

e) Para os próximos 10 clientes que chegam a loja calcule a probabilidade de quatro clientes efetuarem uma compra.

VALOR ESPERADO E VARIÂNCIA DA DISTRIBUIÇÃO BINOMIAL

$$E(X) = \mu = n.p$$
$$Var(x) = \sigma^2 = n.p.(1 - p)$$

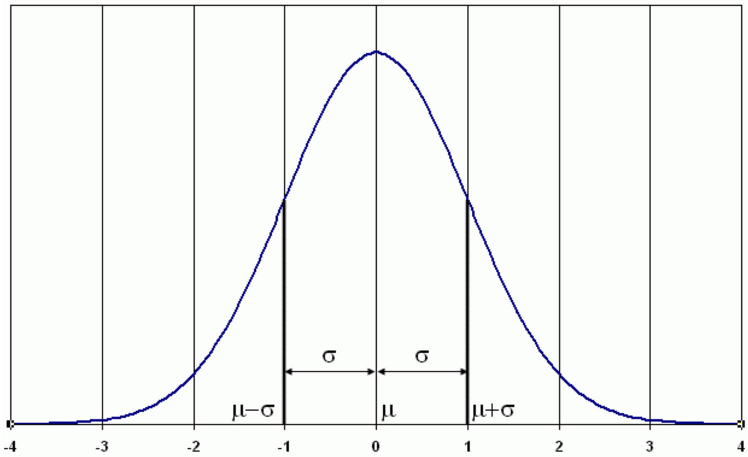
Para o problema 2 calcule:

a) o valor esperado de clientes que farão uma compra; e o desvio padrão;

b) repita os cálculos para o caso de mil clientes entrarem na loja.

DISTRIBUIÇÃO DE PROBABILIDADE CONTÍNUA
DISTRIBUIÇÃO NORMAL

A mais importante distribuição de probabilidade para descrever uma variável contínua é a ***distribuição normal de probabilidade***. Ela é usada em ampla variedade de aplicações práticas em que as variáveis aleatórias são altura, peso, notas de exames, medições científicas, índices pluviométricos, entre outros. É esta distribuição de probabilidade amplamente usada na inferência estatística, razão pela qual vamos estudá-la em detalhes.



A forma da distribuição normal de probabilidades é ilustrada pela curva em forma de sino mostrada na figura. Por envolver uma variável contínua, a função de probabilidade f(x) muda para ***função densidade de probabilidade (fdp)***; a expressão analítica da fdp é dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

onde:

μ	média
σ	desvio padrão
π	3,141592654
e	2,718281828

CARACTERÍSTICAS DA DISTRIBUIÇÃO NORMAL

1. A distribuição normal de probabilidade possui dois parâmetros: a média (μ) e o desvio padrão (σ).
2. O ponto máximo da curva normal encontra-se na média, que é também a mediana e a moda da distribuição.
3. A média da distribuição pode ser qualquer valor numérico: negativo, zero ou positivo.
4. A distribuição normal é simétrica, sendo a forma da curva à

esquerda da média uma imagem espelhada da forma da curva à direita da média.

5. Os extremos da curva (caudas) tendem ao infinito em ambas as direções e, teoricamente, jamais tocam o eixo horizontal.

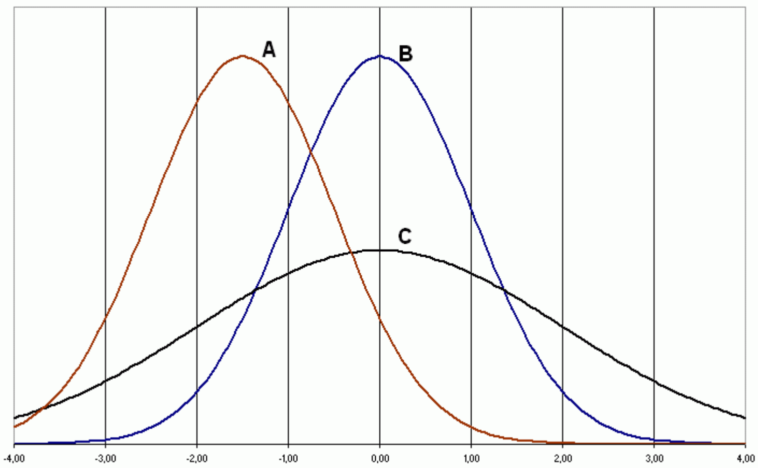
6. O ponto de inflexão da curva mede o valor do desvio padrão; conforme o valor do desvio padrão a curva é achatada ou alongada.

7. Valores maiores do desvio padrão denotam curvas mais largas e mais achatadas, exibindo maior variabilidade dos valores da variável aleatória.

8. As probabilidades da variável aleatória normal são dadas por áreas sob a curva. A área total sob a curva é 1; sendo a distribuição normal simétrica, as áreas à esquerda e à direita da média são iguais a 0,5.

9. As percentagens dos valores de alguns intervalos comumente usados são:

- ✓ 68,3% dos valores de uma variável aleatória normal estão dentro do intervalo: $\mu \pm \sigma$
- ✓ 95,4% dos valores de uma variável aleatória normal estão dentro do intervalo: $\mu \pm 2\sigma$
- ✓ 99,7% dos valores de uma variável aleatória normal estão dentro do intervalo: $\mu \pm 3\sigma$



A figura ilustra três curvas normais; a azul e a vermelha possuem o mesmo desvio padrão, porém apresentam médias diferentes. A outra curva normal (em preto) possui uma forma bastante achatada e alargada, o que evidencia um desvio padrão alto; obviamente há uma grande dispersão de valores.

Para o cálculo das probabilidades de uma variável aleatória normal teríamos que resolver a fdp f(x). Para facilitar, a variável aleatória normal (x) é transformada em uma variável aleatória normal padrão (z) no que resulta a função densidade normal padrão de probabilidade f(z) dada por:

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2}$$

Esta expressão pode ser facilmente calculada e os valores tabelados para diferentes valores de *z*. Tais tabelas fornecem as probabilidades com um único argumento de entrada: *z*.

A normalização ou padronização de uma variável aleatória normal *x* é dada por:

$$z = \frac{x - \mu}{\sigma}$$

A seguir um pequeno trecho da tabela normal padrão: na primeira coluna estão os valores de *z* com uma casa decimal; na primeira linha superior, nas diferentes colunas aparecem valores para a segunda casa decimal de *z*.

Tabela da distribuição normal padrão

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	0,2704	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	0,4306	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	0,4796	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

Aplicações da tabela:

a) Calcular a probabilidade do valor *z* correspondente a variável aleatória normal padrão *x* estar entre 0 e 1.
Relembrando o gráfico, a probabilidade é dada pela área sob a curva compreendida entre os dois extremos, 0 e 1. Assim, o que devemos determinar é P(0 ≤ z ≤ 1).

Na tabela procuramos pelo valor $z = 1,0$ na primeira coluna e na primeira linha procuramos pela coluna representativa das demais decimais de z; como é zero, o valor da probabilidade é dada pela intersecção da linha de $z = 1,0$ e coluna 00; logo o valor procurado é 0,3413.

b) Determinar $P(-1 \leq z \leq 1)$. Como a curva normal é simétrica, o valor da área entre 0 e 1 e entre 0 e -1 é a mesma; vimos em (a) que esta probabilidade é igual a 0,3413. Como o intervalo vai de -1 a 1 a probabilidade deve ser o dobro; assim $P(-1 \leq z \leq 1) = 0,3413 + 0,3413 = 0,6826$.

c) Determinar $P(-2 \leq z \leq 2)$. Da mesma forma a área entre 0 e 2 é 0,4772; assim, pela simetria, $P(-2 \leq z \leq 2) = 0,4772 + 0,4772 = 0,9544$.

d) Determinar $P(-3 \leq z \leq 3)$. Da mesma forma a área entre 0 e 3 é 0,4987; assim, pela simetria, $P(-3 \leq z \leq 3) = 0,4987 + 0,4987 = 0,99974$.

e) Determinar $P(z \geq 1,58)$. O valor coletado na tabela referente a intersecção da linha $z = 1,5$ e a coluna 0,08 é 0,4429. Ora, a curva é simétrica sabemos que 50% da área sob a curva está a direita de $z = 0$ e os outros 50% à esquerda. Assim, queremos a área que está à direita do valor $z = 1,58$; logo, $P(z \geq 1,58) = 0,5 – 0,4429 = 0,0571$.

f) Determinar $P(z \geq -0,5)$. A área entre -0,5 e 0 vale 0,1915. Assim, a $P(z \geq -0,5) = 0,5 + 0,1915 = 0,6915$.

g) Determinar o valor de z tal que a probabilidade de obter um valor z mais elevado seja 0,10.
Agora temos o conhecimento do valor da probabilidade ou da área sob a curva normal; desejamos, portanto, determinar o valor de z que garanta uma área sob a curva (cauda) de 0,10.
Pela simetria a área entre $z = 0$ e z procurado é igual a $0,5 – 0,10 = 0,40$. Fazendo uma varredura no corpo da tabela encontramos 0,3997 como o valor probabilístico mais próximo de 0,4000. Verificando o valor z na coluna extrema esquerda e na linha do topo da tabela encontramos o valor 1,28.

APROXIMAÇÃO NORMAL ÀS PROBABILIDADES BINOMIAIS

Na apresentação da distribuição binomial vimos que o experimento binomial está calcado em dois resultados, apenas: sucesso e fracasso. A variável aleatória binomial é o número de sucessos obtidos nos *n* ensaios e as questões probabilísticas dizem respeito à probabilidade de *x* sucessos nos *n* ensaios.

Quando o número de ensaios torna-se grande é enfadonho o cálculo da função probabilidade binomial, mesmo usando calculadora. Nos casos em que tenhamos: $\boldsymbol{np \geq 5}$ e $\boldsymbol{n(1 - p) \geq 5}$, a distribuição normal fornece uma aproximação fácil de usar às probabilidades binomiais.

Fazemos um ajuste à curva normal calculando os parâmetros μ e σ em função dos parâmetros *n* e *p* da binomial. Assim:

$$\mu = n.p$$
$$\sigma = \sqrt{n.p.(1 - p)}$$

Devemos nos lembrar de que, quando se trata de distribuição contínua de probabilidade, as probabilidades são calculadas como áreas sob a fdp. Consequentemente a probabilidade de um valor único qualquer para a variável aleatória contínua é zero. Corrigimos isso introduzindo o fator de correção de continuidade ***0,5***.

Exemplo: Determinar pela distribuição normal, a probabilidade de $x = 12$ sucessos em 100 ensaios.
Então, a P($x = 12$) da distribuição binomial discreta é aproximada por P($11,5 \leq x \leq 12,5$) da distribuição normal contínua de tal forma que o valor $x = 12$ esteja garantido no intervalo criado.
Na sequência calculamos o valor correspondente de μ e σ e procedemos a normalização para determinar os valores de z.

$\mu = n.p = 100.0,10 = 10$
$\sigma = \sqrt{n.p.(1 - p)} = \sqrt{100.0,10(0,90)} = 3$

Normalizando,
$z_1 = (x - \mu) / \sigma = (11,5 - 10) / 3 = 0,5$
$z_2 = (x - \mu) / \sigma = (12,5 - 10) / 3 = 0,83$

Assim, P($11,5 \leq x \leq 12,5$) = P($0,5 \leq z \leq 0,83$) = 0,2967 – 0,1915 = 0,1052.

Verifique agora, qual a probabilidade de 13 erros ou menos em uma amostra de 100 faturas. (0,8790).

Outros exemplos:

1) Uma distribuição binomial de probabilidade tem p = 0,20 e n = 100.
a) qual é a média e qual é o desvio padrão?
b) é possível fazer a aproximação pela distribuição normal padrão? Explique.
c) qual é a probabilidade de haver exatamente 24 sucessos?
d) qual é a probabilidade de haver entre 18 e 22 sucessos?
e) qual é a probabilidade de haver 15 sucessos ou menos?

2) Uma distribuição binomial de probabilidade tem p = 0,60 e n = 200.
a) qual é a média e qual é o desvio padrão?
b) é possível fazer a aproximação pela distribuição normal padrão? Explique.
c) qual é a probabilidade de haver entre 100 e 110 sucessos?
d) qual é a probabilidade de haver 130 sucessos ou mais?
e) qual é a vantagem de usarmos a distribuição normal de probabilidade para aproximar as probabilidades binomiais? Use o item (d) para explicar a vantagem.

3) Ao assinar um contrato de cartão de crédito você o lê cuidadosamente? Em uma pesquisa para saber "quão minuciosamente você lê um contrato de cartão de crédito", as descobertas revelaram que 44% lêem cada palavra; 33% lêem o suficiente para entender o contrato, 11% dão apenas uma olhada e 4% não o lêem absolutamente.
a) em relação a uma amostra de 500 pessoas, quantas pessoas você acha que diriam que lêem cada palavra de um contrato de cartão de crédito?
b) em relação a uma amostra de 500 pessoas, qual é a probabilidade de 200 ou menos dizerem que lêem cada palavra de um contrato de cartão de crédito?
c) em relação a uma amostra de 500 pessoas, qual é a probabilidade de pelo menos 15 dizerem que não lêem os contratos de cartão de crédito?