Estatística para Metrologia

Aula 4

Mônica Barros, D.Sc.

Abril de 2008

Aula 3

- Variáveis Aleatórias Discretas
  - Bernoulli
  - Binomial
  - Geométrica
  - Binomial Negativa
  - Poisson

Distribuição Bernoulli

- É a mais simples v.a. discreta.
- Seja X uma v.a. com apenas dois valores possíveis, "sucesso" (denotado por 1) e "falha" (denotado por 0).
- Então:

$$f(1) = \Pr(X = 1) = p$$

$$f(0) = \Pr(X = 0) = 1 - p = q$$

- Note que $0 < p < 1$ e "sucesso" e "falha" não indicam se o resultado de uma experiência é "bom"ou "ruim".

Distribuição Bernoulli

- A distribuição de Bernoulli serve como um "tijolo" para a construção de modelos mais elaborados, como a Binomial, a Geométrica e a Binomial Negativa.
- Podemos reescrever a função de probabilidade como:

$$\boxed{f(x) = \Pr(X = x) = p^x (1 - p)^{1-x} \quad \text{onde x} = 0,1}$$

- Esta última notação será útil para identificar uma variável Bernoulli apenas como um caso particular de uma Binomial.
- Notação: X ~ Bernoulli(p)

Distribuição Binomial

- A situação clássica em que usamos uma variável Binomial é:
  - Uma experiência aleatória tem apenas dois resultados possíveis : "sucesso" e "falha", onde a probabilidade de "sucesso" é p e a probabilidade de "falha" é q = 1- p.
- Surge, entre outras aplicações, na amostragem COM reposição.

Distribuição Binomial

- A experiência é repetida um número fixo (n) de vezes, sempre nas mesmas condições, de tal forma que as probabilidades de "sucesso" (p) e "falha" (q =1-p) se mantêm inalteradas a cada repetição.
- As diversas repetições da experiência são feitas de maneira independente, ou seja, o resultado de uma repetição não afeta o resultado das outras.

Distribuição Binomial

- A variável aleatória X que mede o número de "sucessos" nas n repetições da experiência é uma variável discreta, com valores possíveis 0, 1, 2, ... , n.
- Dizemos que esta variável tem distribuição Binomial com parâmetros n e p, e escrevemos X ~ Bin(n, p).

Distribuição Binomial - quadro resumo

experiência aleatória	"sucesso"	"falha"	p = probabilidade de "sucesso"	n = número de repetições da experiência	X = variável aleatória Binomial
"chutar" a resposta numa prova de múltipla escolha onde cada questão tem 5 opções	acertar a resposta da questão	errar a resposta	1/5	número de questões da prova	número de respostas certas na prova
nascimento de uma criança numa família	menina	menino	1/2	número de crianças na família	número de meninas na família
jogada de um dado	sair o número 6	sair qualquer outro número	1/6	número de jogadas do dado	número de vezes em que saiu o número 6 nas n jogadas do dado
verificar se uma peça produzida numa fábrica tem defeito	peça tem defeito	peça não tem defeito	proporção de peças com defeito na população de peças	tamanho da amostra	número de peças com defeito na amostra

Distribuição Binomial – exemplo típico

- É importante ter em mente um exemplo típico, um modelo da situação que representa a aplicação de uma distribuição de uma densidade ou função de probabilidade. No caso da Binomial eu sempre sugiro a idéia da nota da prova de múltipla escolha em chinês ....
- X = número de questões “chutadas” certo numa prova de múltipla escolha em chinês (presumindo que você não saiba chinês!). Todas as questões têm a mesma probabilidade de acerto e o fato de você acertar ou errar uma questão não afeta a probabilidade das outras.
- X é a sua nota na prova.

Distribuição Binomial

- Se X ~ Bin(n,p), sua função de probabilidade é:

$$f(x) = \Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \quad \text{para } x = 0,1,2,\ldots,n$$

- No Excel:
  - Use a função estatística distrbinom (ou dist.binom, dependendo da versão do Excel).

Distribuição Binomial

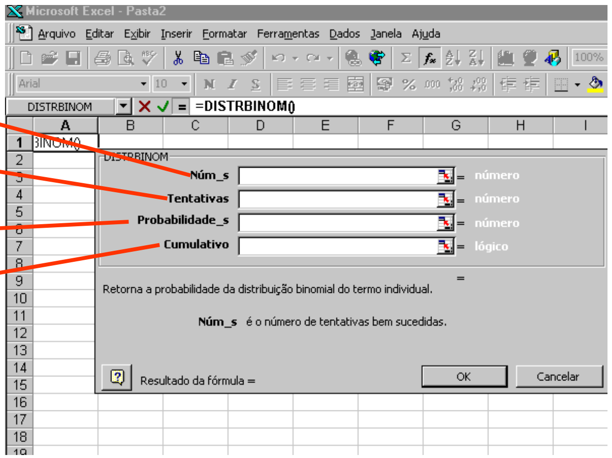
- No Excel:

Valor de X

Valor de n

Valor de p

Argumento lógico - se VERDADEIRO produz a função de distribuição acumulada, se FALSO produz a função de probabilidade



Distribuição Binomial

- Média e Variância
- Se X ~ Bin(n,p) então:
  - E(X) = n.p
  - VAR(X) = n.p.q = n.p.(1-p)
- Nota: a distribuição de Bernoulli é apenas um caso particular da Binomial com n = 1. Logo, segue que a média e a variância de uma variável Bernoulli(p) são, respectivamente, p e p.q.

Distribuição Binomial

- Exemplo
- Em uma loja, a probabilidade de um cliente realizar uma compra é de 15%. Qual a probabilidade de, entre 5 clientes que entram na loja, exatamente 3 realizarem uma compra?

$$n = 5; \quad x = 3; \quad p = 0.15$$

$$f(3) = \frac{5!}{3!.2!}(0.15)^3.(0.85)^2 = 0.024$$

Distribuição Binomial

- Exemplo

  Numa eleição supõe-se que 30% dos eleitores são favoráveis a uma certa proposta. Toma-se uma amostra de tamanho 20 de eleitores da cidade do Rio de Janeiro. Calcular as probabilidades de 4, 5, 6, 7, 8, 9 ou 10 dos eleitores na amostra serem favoráveis à proposta.
- Solução

  Seja X o número de eleitores na amostra que são a favor da proposta. Então os valores possíveis de X são 0, 1, 2, ..., 20, e X tem distribuição Binomial com parâmetros n = 20 e p = 0.30.

Distribuição Binomial

- As probabilidades para os diversos valores de X são calculadas através da fórmula:

$$\Pr(X = x) = \binom{20}{x}.(0.30)^x.(0.70)^{20-x} = \frac{20!}{x!(20-x)!}.(0.30)^x.(0.70)^{20-x}$$

- A tabela a seguir foi produzida usando a função distrbinom do Excel para x = 4, 5, ..., 10.

x	Pr(X =x)
4	13.04%
5	17.89%
6	19.16%
7	16.43%
8	11.44%
9	6.50%
10	3.08%

Exemplo (para casa – use o Excel)

- Uma empresa aérea sabe que 20% das pessoas que fazem reservas aéreas cancelam suas reservas.
- A empresa vende 50 passagens para um vôo que contém apenas 46 lugares. Supondo que as pessoas cancelam suas reservas de maneira independente, calcule a probabilidade de que haverá assentos para todos os passageiros.

Exemplo (para casa – use o Excel)

- Você arranjou um emprego numa pizzaria que funciona no sistema de entrega a domicílio. Apenas 5% dos pedidos são de pizza de lombinho com abacaxi.

  Você recebe exatamente 9 pedidos pelo telefone, qual a probabilidade de, no máximo, 1 pizza de lombinho com abacaxi ser pedida?

  Você recebe exatamente 30 pedidos pelo telefone num dia de bastante movimento, qual a probabilidade de receber mais de 3 pedidos de pizza de lombinho com abacaxi? Dica: pare e pense antes de fazer contas desnecessárias!!!!!!

Distribuição Geométrica

- Da mesma maneira que a Distribuição Binomial, a Geométrica também parte de repetições de Bernoulli independentes.
- A grande diferença em relação à Binomial é que, no caso da Geométrica, *o número de repetições não é fixo*, e as repetições são feitas até encontrar o primeiro "sucesso" (e então a experiência pára).

Distribuição Geométrica

- Um exemplo típico é: imagine uma caixa com um número muito grande de peças. Estas peças podem estar OK ou defeituososas, e a probabilidade de encontrar uma peça defeituosa é FIXA e igual a "p".
- Retiram-se peças da caixa até encontrar uma peça com defeito.

Distribuição Geométrica

- Neste exemplo, um "sucesso" indica encontrar uma peça defeituosa.
- Note que estamos supondo que a caixa onde estão as peças contém um número muito grande de objetos, e assim a probabilidade de retirada de uma peça defeituosa é mantida constante (p).

Distribuição Geométrica

- Seja X o número de repetições necessárias até encontrar uma peça com defeito.
- Então a função de probabilidade de X é:

$$f(x) = \Pr(X = x) = q^{x-1}.p = (1-p)^{x-1} p \quad \text{onde } x = 1,2,3,....$$

- Pode-se provar, usando a série geométrica, que f(x) assim definida é uma função de probabilidade.

Distribuição Geométrica - Exemplo

- Você arranjou um emprego numa empresa que faz pesquisas de opinião pelo telefone.
- Apenas 10% das chamadas resultam numa pesquisa completa, isto é, apenas 10% dos entrevistados responde todo o seu questionário. Calcule as seguintes probabilidades:
- a) De que a primeira pesquisa completa será respondida na 5a. ligação telefônica.
- b) De que a primeira pesquisa completa será respondida na 8a. ligação telefônica.

Distribuição Geométrica - Exemplo

- Esta é uma aplicação típica da distribuição Geométrica.
- Seja X o número de ligações efetuadas até que a 1a. pesquisa completa seja respondida.
- Então X é uma v.a. Geométrica com probabilidade de sucesso p = 0.10.
- a) $\Pr(X = 5) = (0.90)^4 (0.10) = 6.56\%$
- b) $\Pr(X = 8) = (0.90)^7 (0.10) = 4.78\%$

Distribuição Geométrica – para casa

- Exemplo
- Um gestor de fundos de investimento ultrapassa a sua meta de retorno mensal 85% das vezes e nos 15% restantes tem um resultado ruim (abaixo da meta).
- Qual a probabilidade dele ter o primeiro resultado ruim nos próximos 12 meses?
- E nos primeiros 6 meses?

Distribuição Geométrica – para casa

- Todo final de semana você vai para a sua casa de campo. Você é meio apressado e gosta de ultrapassar o limite de velocidade na estrada. A probabilidade do radar pegar você acima da velocidade permitida é 15%.
- Se você é pego pela polícia tem que pagar uma multa de R$ 250,00 (por que, além de tudo você sempre esquece os documentos do carro em casa ....).

Distribuição Geométrica – para casa

- Suponha que cada ida para o campo no fim de semana seja uma repetição independente. O custo associado a cada viagem é R$ 25,00 (gasolina e pedágio).
- Você continua dirigindo em alta velocidade até receber a primeira multa.
- a) Qual o custo esperado deste procedimento (viajar em alta velocidade até ganhar a primeira multa)?
- b) Suponha que você tenha disponível R$ 300,00 no banco. Qual a probabilidade de você estourar o seu orçamento com este procedimento?

Distribuição Geométrica

- A função de probabilidade de X é:

$$f(x) = \Pr(X = x) = q^{x-1}.p \quad \text{onde } x = 1, 2, 3, ...$$

- Notação: X ~ Geom(p)
- Média e Variância
- Se X ~ Geom(p) então:
  1) E(X) = 1/p
  2) VAR(X) = q/p^2

Distribuição Binomial Negativa

- A distribuição Binomial negativa é mais uma função de probabilidade derivadas de tentativas de Bernoulli independentes e é uma ***generalização da distribuição Geométrica***.
- Suponha que *repetimos um número indefinido de vezes* uma experiência que resulta em sucesso ou falha. As repetições terminam quando encontramos o r-ésimo sucesso, onde r é um número especificado a priori.

Distribuição Binomial Negativa

- Seja X a variável aleatória que representa a tentativa onde o r-ésimo sucesso ocorre. Então a função de probabilidade de X é:

$$f(x) = \Pr(X = x) = \binom{x-1}{r-1} . p^r . q^{x-r} \quad \text{onde } x = r, r+1, r+2, ...$$

- Note que o valor mínimo de X é r, pois precisamos fazer pelo menos r repetições para encontrar r sucessos! Também, a combinação que aparece na densidade indica que, das x-1 repetições anteriores à última (que é necessariamente um "sucesso", r-1 são "sucessos").

Distribuição Binomial Negativa

- Também, $p^r.q^{x-r}$ é a probabilidade de uma seqüência qualquer contendo r "sucessos" e x-r "falhas".
- Se X ~ NegBin(r, p), sua média e variância são:

E(X) = r/p

VAR(X) = rq/p²

- Nota – a função de probabilidade Binomial Negativa com parâmetros r = 1 e p é apenas a Geométrica.
- Na verdade, uma variável Binomial Negativa (r, p) é apenas a soma de r Geométricas(p), todas independentes. Faz sentido, não?

Distribuição Binomial Negativa - Exemplo

- Um comprador em potencial entra numa loja de carros a cada hora.
- Um vendedor tem probabilidade 0.25 de concluir uma venda. O vendedor decide trabalhar até conseguir vender 3 carros num só dia.
- Qual a probabilidade de que o vendedor tenha de trabalhar exatamente 8 horas para conseguir vender os 3 carros? E mais de 8 horas?

Distribuição Binomial Negativa - Exemplo

- Seja X o número de horas de trabalho necessárias para vender 3 carros.
- Então X tem distribuição Binomial Negativa com parâmetros r = 3 e p = 0.25.

$$\Pr(X = x) = \binom{x-1}{2}(0.25)^3.(0.75)^{x-3} \quad \text{para } x = 3, 4, 5, ...$$

- A próxima tabela exibe os valores das probabilidades para X = 3, 4, ..., 8.

Distribuição Binomial Negativa - Exemplo

x	Pr(X = x)
3	0.01563
4	0.03516
5	0.05273
6	0.06592
7	0.07416
8	0.07787

A probabilidade de trabalhar exatamente 8 horas é Pr (X = 8) = 0.07787. A probabilidade de trabalhar mais de 8 horas é:

$$\Pr(X > 8) = 1 - \Pr(X \le 8) = 1 - \{\Pr(X = 3) + \Pr(X = 4) + ... + \Pr(X = 8)\} =$$
$$= 1 - \{0.01563 + 0.03516 + ... + 0.07787\} = 1 - 0.32146 = 0.67854$$

Distribuição Binomial Negativa - Exemplo

- **Uma fábrica de sorvetes decidiu fazer uma campanha para aumentar suas vendas.**
- **A cada 50 sorvetes produzidos um é premiado, e o prêmio consiste em ganhar um outro sorvete grátis. Cada sorvete é vendido por R$ 0.80.**
- **a) Se você decide comprar sorvetes até encontrar um sorvete premiado, quanto você espera gastar?**
- **b) E se você comprar sorvetes até encontrar o 2o. sorvete premiado?**

Distribuição Binomial Negativa - Exemplo

- **a) Seja X o número de sorvetes comprados até encontrar o 1o. sorvete premiado. Então X tem distribuição Geométrica com parâmetro p =1/50. Seja C o custo deste procedimento. Então C = 0.8X e E(C) = 0.8.E(X). Mas, E(X) = 1/p = 50 e assim o custo esperado é de R$40.**

- **b) Neste caso X mede o número de sorvetes comprados até encontrar o 2o. sorvete premiado. Então X tem distribuição Binomial Negativa com r = 2 e p =1/50. Agora E(C) = 0.8E(X) = 0.8(2)(50) = R$ 80.**

Distribuição Binomial Negativa – para casa

- **Uma gulosa professora de estatística é "fissurada" por trufas de chocolate. Em busca da trufa ideal, ela vai provando chocolates em diversas lojas, de maneira independente.**

- **A probabilidade dela gostar de uma trufa que prova é 70%. Ela decide passear por um shopping, provando todas as trufas que encontra, e decide parar só ao encontrar a 4ª. trufa "maravilhosa"(para "desespero" da balança que tem em casa!).**

Distribuição Binomial Negativa – para casa

- Qual a probabilidade dela ter que:
- Provar 6 trufas até encontrar a 4ª. trufa maravilhosa?
- Ter que "sofrer", provando 10 trufas, até encontrar a 4ª. trufa maravilhosa?

Distribuição de Poisson

- Está associada a experiências que modelam o número de ocorrências de um evento dentro de um determinado intervalo de tempo (ou espaço), quando estes eventos ocorrem com uma taxa média conhecida, por exemplo:
  - Número de carros que passam por uma estrada no intervalo de uma hora
  - Número de buracos por km de uma rodovia
  - Número de assassinatos num final de semana
  - Número de defeitos por metro de tecido produzido

Distribuição de Poisson

- Número de erros de digitação numa página de texto
- Número de mutações num trecho de DNA após a exposição a uma certa quantidade de radiação
- Número de soldados mortos por chutes de cavalo a cada ano na cavalaria Prussiana. Este exemplo ficou famoso num livro de Bortkiewicz (1868–1931).

Distribuição de Poisson

- Um experimento de Poisson possui as seguintes características:
  - A probabilidade de uma ocorrência é a mesma para intervalos iguais;
  - As ocorrências são independentes, e independente do tempo em que o último evento aconteceu.
- A distribuição foi descoberta por Siméon-Denis Poisson (1781–1840) e publicada em 1838 em seu trabalho "Recherches sur la probabilité des jugements en matières criminelles et matière civile" (Wikipedia).

Distribuição de Poisson

- A função de probabilidade para uma variável Poisson com parâmetro μ é:

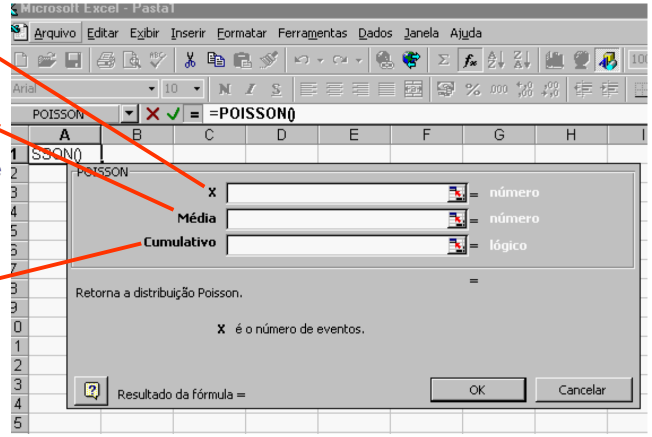
$$f(x) = \Pr(X = x) = \frac{\mu^x e^{-\mu}}{x!} \quad \text{onde} \quad x = 0,1,2,....$$

- Este parâmetro μ é a MÉDIA da distribuição e indica o número esperado de ocorrências num dado intervalo.
- A distribuição Poisson é freqüentemente utilizada na modelagem de eventos "raros", ou seja, a probabilidade de X = 0 ou X = 1 (pequeno número de ocorrências no intervalo de tempo especificado) é grande.

Distribuição de Poisson no Excel

Valor de X

Parâmetro da função de probabilidade



Argumento lógico

Se FALSO fornece a função de probabilidade f(x)

Distribuição de Poisson

- Média e Variância
- Se X ~ Poisson(μ) então:
  - E(X) = μ
  - VAR(X) = μ

Distribuição de Poisson

- Exemplo
- O número médio de clientes que entram em um banco num período de 15 minutos é 10.
- Qual a probabilidade de entrarem exatamente 5 clientes em 15 minutos?

$$\mu = 10; \quad x = 5$$

$$f(5) = \Pr(X = 5) = \frac{10^5 e^{-10}}{5!} = 0.0378$$

Distribuição de Poisson

- **Exemplo**
- **Numa campanha de caridade feita por um programa de TV em todo o Brasil, o número de pessoas que contribuem mais de 500 reais é uma variável aleatória com média de 5 pessoas por programa.**
  - **a) Calcule a probabilidade de que, num certo programa, o número de pessoas que contribuem mais de 500 reais exceda 8.**
  - **b) Faça o gráfico da função de probabilidade.**
  - **c) Faça um gráfico da função de distribuição acumulada.**

Distribuição de Poisson

- **Solução**

Seja X o número de pessoas que contribuem com mais de 500 reais a cada programa. Desejamos calcular Pr{ X > 8}.

$\Pr\{X > 8\} = 1 - \Pr\{X \leq 8\} = 1 - F(8)$

onde F(.) denota a função de distribuição acumulada.

A tabela a seguir apresenta a função de probabilidade, a função de distribuição acumulada e seu complemento.

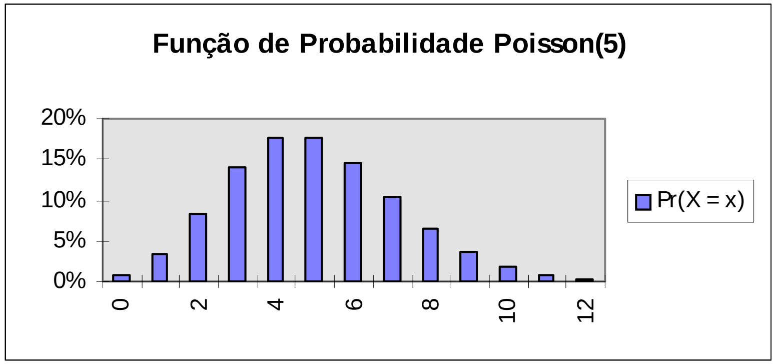
Da tabela segue que Pr(X > 8) = 6.81%.

Distribuição de Poisson

x	Pr(X = x)	Pr(X <= x)	1- F(x) = Pr(X > x)
0	0.67%	0.67%	99.33%
1	3.37%	4.04%	95.96%
2	8.42%	12.47%	87.53%
3	14.04%	26.50%	73.50%
4	17.55%	44.05%	55.95%
5	17.55%	61.60%	38.40%
6	14.62%	76.22%	23.78%
7	10.44%	86.66%	13.34%
8	**6.53%**	**93.19%**	**6.81%**
9	3.63%	96.82%	3.18%
10	1.81%	98.63%	1.37%
11	0.82%	99.45%	0.55%
12	0.34%	99.80%	0.20%

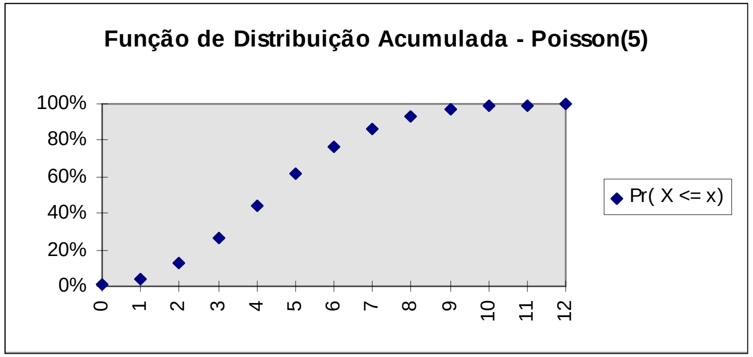
Distribuição de Poisson

b) A função de probabilidade é dada no gráfico abaixo



Distribuição de Poisson

c) A função de distribuição acumulada é mostrada a seguir.



Distribuição de Poisson

- O número de chamadas para um telefone com prefixo 800 (chamadas grátis) é uma variável aleatória com média de 3 chamadas por minuto.
- Qual a probabilidade do número de chamadas num minuto ser maior que 4?
- Solução
- Suponha que a distribuição do número de chamadas é Poisson com a média indicada (3 chamadas por minuto).

Distrbuição de Poisson

- Logo, a função de probabilidade é:

$$f(x) = \Pr(X = x) = \frac{e^{-3}(3)^x}{x!} \quad \text{para } x = 0,1,2,....$$

- A probabilidade desejada é:

$$\Pr(X > 4) = \Pr(X = 5) + \Pr(X = 6) + ... = 1 - \Pr(X \le 4) = 1 - \sum_{x=0}^{4} \frac{e^{-3}(3)^x}{x!} =$$

$$= 1 - e^{-3}\left\{1 + 3 + \frac{9}{2} + \frac{27}{6} + \frac{81}{24}\right\} = 81.53\%$$

Exemplo (para casa – use o Excel)

- O número de enchentes em cada verão no Rio de Janeiro é uma variável aleatória Poisson com média de 2 enchentes por verão.
  - Calcule a probabilidade de ocorrerem exatamente 3 enchentes em um verão qualquer.
  - Calcule a probabilidade de ocorrerem menos de 10 enchentes em 30 verões.

Exemplo (para casa – use o Excel)

- O número de carros que chegam num posto de pedágio é uma variável Poisson com parâmetro 3 carros por minuto.
- Use o Excel para calcular:
  - A probabilidade de passarem mais de 4 carros num minuto.
  - A probabilidade de passarem menos de 25 carros em 10 minutos.

Exemplo (para casa – use o Excel)

- O número médio de pedidos de autorização para um certo exame médico complexo recebido por um plano de saúde é uma variável Poisson com parâmetro λ = 4 pedidos por hora.
- Calcule a probabilidade de, numa hora qualquer, a empresa receber mais de 5 pedidos de autorização para este exame.
- Calcule a probabilidade da empresa recebem, em uma hora, 9 ou menos pedidos de autorização.

Distribuição Poisson – para casa

- O número de erros de digitação numa página de livro é uma variável aleatória Poisson com média de 2 erros por página. Um capítulo contém 30 páginas. Calcule as seguintes probabilidades:
- a) De que o número total de erros seja menor que 12.
- b) De que o número total de erros exceda 10.