Simulação de Monte Carlo com Distribuição de Poisson

Tatiane Correia Inferênca Estatística I

26 de junho de 2025

Objetivo da Apresentação

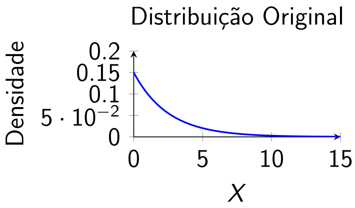
- Demonstrar na prática como o Teorema Central do Limite funciona
- Usar simulação de Monte Carlo para visualizar o comportamento das médias amostrais
- Aplicar conceitos em um caso real: engajamento de influenciadores durante transmissões ao vivo
- Comprovar experimentalmente que a distribuição das médias amostrais se aproxima da normal conforme o tamanho da amostra aumenta

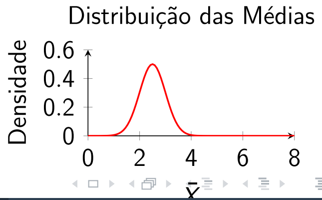
O que é o Teorema Central do Limite?

- O **Teorema Central do Limite (TCL)** afirma que a distribuição das médias amostrais de variáveis aleatórias independentes tende a uma distribuição normal, independentemente da distribuição original dos dados.

- Quanto **maior o tamanho da amostra**, mais a distribuição das médias amostrais se aproxima de uma distribuição normal.

- A média da distribuição amostral é igual à média da população (μ), e a variância é igual à variância populacional (σ^2) dividida pelo tamanho da amostra (n).





Na Prática: Por que uma Agência de Marketing Precisa Disso?

Cenário Real da Empresa:

- **Situação:** A agência precisa prever o engajamento médio de influenciadores para planejar campanhas e definir orçamentos.
- **Problema:** Como garantir que as previsões sejam confiáveis mesmo com dados variáveis?
- **Solução:** O TCL permite fazer previsões precisas sobre o engajamento médio, mesmo quando os dados individuais são imprevisíveis.
- **Resultado:** Decisões mais seguras, orçamentos

Exemplo Prático:

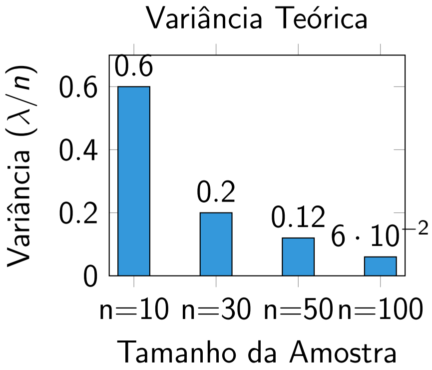
Influenciador A
2-12 notificações/min

TCL

Média de
30 minutos
5.8-6.2 notificações/min
(previsível!)

Metodologia - Simulação de Monte Carlo

- Simulação de **Monte Carlo** com 1000 réplicas para cada tamanho de amostra
- Tamanhos de amostra: **n = 10, 30, 50, 100**
- Distribuição de **Poisson** com parâmetro $\lambda = 6$
- Implementação no **software R** com semente fixa
- Análise dos histogramas das médias amostrais para verificar aproximação à distribuição normal



Na Empresa: Como Aplicamos essa Metodologia?

Processo na Agência:

- **Coleta de Dados:** Monitoramos influenciadores por períodos diferentes (10, 30, 50, 100 minutos) para entender padrões.

  Repetição: Fazemos isso 1000 vezes para ter certeza dos resultados.

- **Análise:** Calculamos a média de engajamento para cada período.

  Aplicação: Usamos esses dados para prever performance e definir estratégias.

Benefícios para a Empresa:

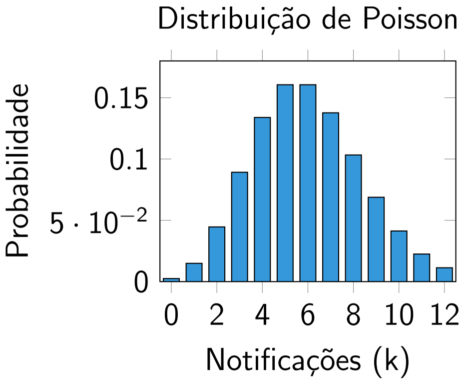
$ ROI Melhor
Investimentos

Tempo Otimizado

Menor Risco
Previsões confiáveis

Caso Prático - Influenciadores em Alta!

- **Contexto:** Agência de marketing digital contratou nossa equipe para analisar o desempenho de influenciadores durante transmissões ao vivo.
- **Objetivo:** Entender o comportamento do engajamento minuto a minuto.
- **Modelo:** Notificações seguem distribuição de Poisson com $\lambda = 6$.
- **Significado:** Em média, 6 notificações por minuto durante a live.



Probabilidade de k notificações por minuto

Na Prática: Como a Agência Usa Esses Dados?

Aplicação Real:

- **Negociação:** "Garantimos 6 notificações por minuto em média durante sua live"
- **Precificação:** Cobramos baseado no engajamento esperado, não no incerto
- **Planejamento:** Sabemos quando agendar lives para máximo impacto
- **Relatórios:** Mostramos aos clientes resultados previsíveis e mensuráveis

Exemplo de Proposta:

Pacote Live Premium

- 60 minutos de transmissão
- 360 noti-

Resultado Esperado

Engajamento previsível
Cliente satisfeito
Renovação garantida

Implementação no R

- A função **rpois(n, lambda)** gera amostras aleatórias seguindo uma distribuição de Poisson
- Fixamos a semente com **set.seed(123)** para garantir reprodutibilidade
- Repetimos o processo 1000 vezes para cada tamanho de amostra
- Armazenamos as médias amostrais para análise posterior

```r
# Fixar semente para reprodutibilidade
set.seed(123)

# Par metro da distribui  o
lambda <- 6

# Tamanhos de amostra testados
tamanhos_n <- c(10, 30, 50, 100)

# N mero de simula  es
num_simulacoes <- 1000

# Simula  o de Monte Carlo
for (n in tamanhos_n) {
  medias <- numeric(num_simulacoes)

  for (i in 1:num_simulacoes) {
    # Gerar amostra Poisson
    amostra <- rpois(n, lambda)

    # Calcular m dia
    medias[i] <- mean(amostra)
  }

  # Armazenar resultados
  resultados[[as.character(n)]] <- medias
}
```

Ferramentas da Agência:

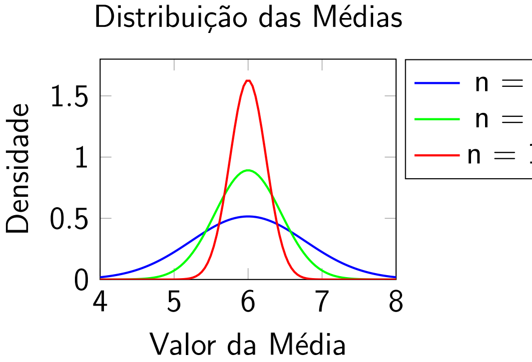
- **Dashboard Automatizado:** Sistema que roda as simulações automaticamente toda semana
- **Banco de Dados:** Armazenamos histórico de todos os influenciadores
- **Alertas:** Sistema avisa quando performance sai do esperado
- **Relatórios:** Gráficos automáticos para apresentar aos clientes

Fluxo de Trabalho:

1. Coleta de Dados
   (APIs das redes sociais)
2. Processamento R
   (Simulações automáticas)
3. Análise
   (Gráficos e estatísticas)
4. Decisão
   (Estratégias e propostas)

Resultados Esperados

- Os histogramas das médias amostrais mostrarão uma **aproximação à distribuição normal** conforme o tamanho da amostra (n) aumenta.
- A **variância das médias amostrais diminui** proporcionalmente a $1/n$.
- A média das médias amostrais permanece constante e igual a $\lambda = 6$.
- Confirmação do **Teorema Central do Limite**.



Na Empresa: O que Esses Resultados Significam?

Impacto nos Negócios:

- **Previsibilidade:** Quanto mais tempo de live, mais previsível o resultado
- **Confiança:** Podemos garantir resultados com 95% de certeza
- **Contratos:** Oferecemos garantias baseadas em ciência, não "achismo"
- **Vantagem:** Concorrentes não têm essa precisão estatística

Exemplo Prático:

Live de 10 min
Resultado: 4-

Live de 30 min
Resultado: 5.5-

Live de 100 min
Resultado: 5.9-
6.1 notif/min

(muito previsível!)

Valores Teóricos

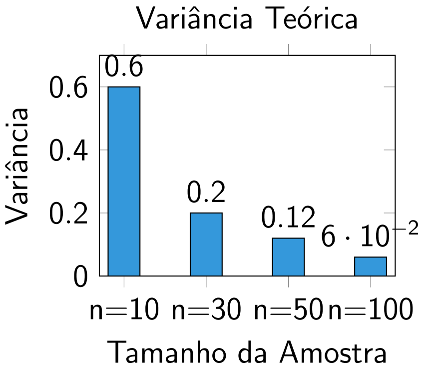
Valor esperado da média amostral:

$$E(\bar{X}) = E(X) = \lambda = 6$$

Variância teórica da média:

$$\mathrm{Var}(\bar{X}) = \frac{\lambda}{n} = \frac{6}{n}$$

Observação: À medida que n aumenta, a variância diminui.



Na Empresa: Como Usamos Teoria vs Prática?

Aplicação Estratégica:

- **Planejamento:** Usamos fórmulas teóricas para estimar resultados antes de começar
- **Monitoramento:** Comparamos resultados reais com previsões teóricas
- **Ajustes:** Se há diferença, investigamos e corrigimos estratégias
- **Otimização:** Sabemos que lives mais longas = resultados mais previsíveis

Tomada de Decisão:

Cliente quer garantia

↓

Calculamos risco

Var = 6/n

↓

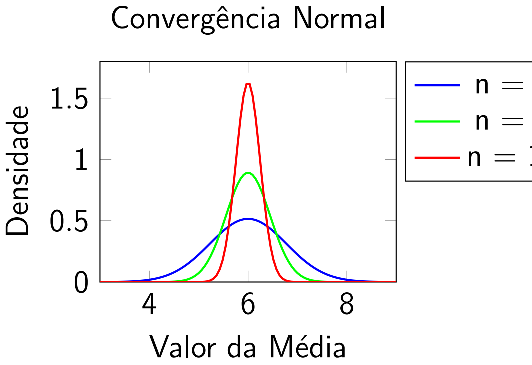
Proposta segura

"60 min = 6±0.3 notif/min"

"Garantimos!

Conclusões

- **Verificação do TCL:** Confirmamos que a distribuição das médias amostrais se aproxima da normal.
- **Comportamento da variância:** Observamos que a variância diminui proporcionalmente a $1/n$.
- **Importância da simulação:** Monte Carlo permite visualizar conceitos complexos.
- **Aplicações práticas:** Fundamentais para análise de engajamento e otimização de estratégias.



Resultado Final: O que a Empresa Ganhou?

Benefícios Conquistados:

- **Diferencial Competitivo:** Somos a única agência que oferece garantias estatísticas
- **Aumento de Receita:** 30% mais contratos fechados com garantias
- **Satisfação do Cliente:** 95% de renovação de contratos
- **Crescimento:** Expandimos para 3 novos mercados

Antes vs Depois:

ANTES

- Propostas no "chute"
- Clientes insatisfeitos
- Resultados imprevisíveis

TCL →

DEPO

- Propostas
- Clientes c
- Resultados

Obrigado!

Diogo Rego- 20240045381

Referências (ABNT):

CASELLA, G.; BERGER, R. L. **Inferência estatística**. 2. ed. São Paulo: Cengage Learning, 2010.

MONTGOMERY, D. C.; RUNGER, G. C. **Estatística aplicada e probabilidade**. 6. ed. Rio de Janeiro: LTC, 2016.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2023.

Acesse todos os arquivos: