Probabilidade e Estatística - Distribuições Gama, Exponencial e χ^2

Distribuição Gama

Usada quando se sabe que as variáveis aleatórias são sempre positivas. Tem como um caso particular as distribuições exponeciais. São relacionadas com eventos de Poisson ou com processos que não tem memória.

Exemplo: Consideremos o caso em que uma empresa constata que a f.d.p. do tempo de vida x de uma lâmpada produzida por ela é dada por:

$$f(x) = \begin{cases} \exp(-x), & \text{para } x > 0 \\ 0, & \text{caso contrário} \end{cases}$$

Caso a empresa queira computar o tempo de vida médio e a variância do tempo de vida então teria que computar:

$$\mathsf{E}[X] = \int_0^\infty x\, e^{-x}\, dx$$

e

$$\mathsf{E}\left[X^2\right] = \int_0^\infty x^2\, e^{-x}\, dx$$

Para computar estas integrais podemos utilizar mão da chamada função gama Γ.

Função Gama

Definição 1 (Função Gama $\Gamma(\alpha)$): *Para cada α positivo, a função gama $\Gamma(\alpha)$ é definida como:*

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}\, e^{-x}\, dx$$

A partir da definição temos:

$$\Gamma(1) = \int_0^\infty e^{-x}\, dx = -e^{-x}\big|_0^\infty = 1$$

Teorema 1: *Seja $\alpha > 1$, então:*

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$$

Isto pode ser demonstrado se realizarmos a integração de $\Gamma(\alpha)$ por partes. Neste caso podemos definir $u = x^{\alpha-1}$, que resulta em $du = (\alpha-1)x^{\alpha-2}dx$. Também podemos definir $dv = e^{-x}dx$, que resulta em $v = -e^{-x}$. Desta forma:

$$\begin{aligned}\Gamma(\alpha) &= -\left[x^{\alpha-1}e^{-x}\right]_0^\infty + \int_0^\infty (\alpha-1)x^{\alpha-2}e^{-x}\, dx \\ &= 0 + (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}\, dx \\ &= (\alpha-1)\Gamma(\alpha-1)\end{aligned}$$

Por conta da propriedade acima, podemos ver que se n é um inteiro positivo, então:

$$\begin{aligned}\Gamma(\alpha) &= (\alpha-1)\Gamma(\alpha-1) \\ &= (\alpha-1)(\alpha-2)\Gamma(\alpha-2) \\ &= (\alpha-1)(\alpha-2)(\alpha-3)\Gamma(\alpha-3) \\ &= (\alpha-1)(\alpha-2)...(1)\Gamma(1) \\ &= (\alpha-1)!\end{aligned}$$

Voltando ao Exemplo: Agora podemos então calcular as integrais buscadas. Podemos ver que:

$$\begin{aligned}\mathsf{E}[X] &= \int_0^\infty x\, e^{-x}\, dx \\ &= \int_0^\infty x^{2-1}\, e^{-x}\, dx \\ &= \Gamma(2) \\ &= 1\end{aligned}$$

Adicionalmente:

$$\begin{aligned}\mathsf{E}[X] &= \int_0^\infty x^2\, e^{-x}\, dx \\ &= \int_0^\infty x^{3-1}\, e^{-x}\, dx \\ &= \Gamma(3) \\ &= 2\end{aligned}$$

Assim, o valor médio do tempo de vida das lâmpadas é 1, e a variância buscada é $\sigma^2 = \mathsf{E}\left[X^2\right] - \mathsf{E}[X] = 2 - 1^2 = 1$.

Em muitas aplicações, é importante de se calcular o valor de $\Gamma(n+1/2)$, com n um inteiro positivo. Nestes casos temos:

$$\Gamma(n+1/2) = (n-3/2)(n-5/2)...(1/2)\Gamma(1/2).$$

Teorema 2: *O valor de $\Gamma(\frac{1}{2})$ é $\sqrt{\pi}$.*

Prova: Partindo-se da definição de $\Gamma(\alpha)$ temos:

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty x^{-1/2}\, e^{-x}\, dx.$$

Podemos fazer a seguinte substituição: $x = (1/2)y^2$, de forma que temos $x^{-1/2} = (1/2)^{-1/2}y^{-1}$ e $dx = ydy$, que resulta em

$$\begin{aligned}\Gamma(\tfrac{1}{2}) &= \int_0^\infty 2^{1/2}y^{-1}\, e^{-(1/2)y^2}\, y\, dy \\ &= 2^{1/2}\int_0^\infty e^{-(1/2)y^2}\, dy \qquad (1)\end{aligned}$$

A integral na equação acima pode ser encontrada se lembrarmos da normalização da distribuição normal, com $\mu = 0$ e $\sigma = 1$:

$$\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}x^2\right]\, dx = 1$$

que resulta em:

$$\int_{-\infty}^\infty \exp\left[-\frac{1}{2}x^2\right]\, dx = \sqrt{2\pi}$$

Como a função integrada é par, ou seja, é simétrica ao redor de $x = 0$, então temos que:

$$\int_0^\infty \exp\left[-\frac{1}{2}x^2\right] dx = \frac{\sqrt{2\pi}}{2}.$$

Substituindo-se esta equação na equação (1) temos:

$$\Gamma(\frac{1}{2}) = 2^{1/2}\frac{\sqrt{2\pi}}{2} = \sqrt{\pi}$$

Assim, voltando ao problema inicial, temos:

$$\Gamma(n + 1/2) = (n - 1/2)(n - 3/2)...(1/2)\sqrt{\pi}.$$

Teorema 3: *Para $\alpha > 0$ e $\beta > 0$ temos:*

$$\int_0^\infty x^{\alpha-1}e^{-\beta x}\, dx = \frac{\Gamma(\alpha)}{\beta^\alpha}. \qquad (2)$$

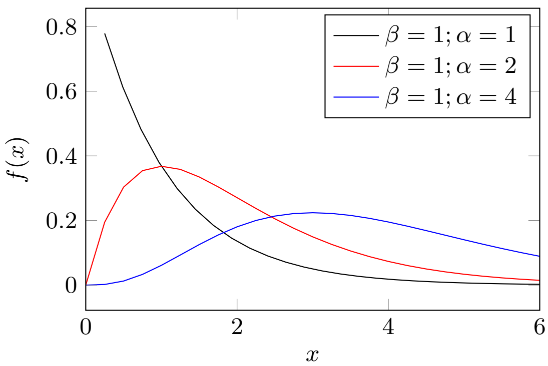
Prova: Se fizermos a substituição de $y = \beta x$, temos $x = y/\beta$ e $dx = dy/\beta$. Assim:

$$\begin{aligned}\int_0^\infty x^{\alpha-1}e^{-\beta x}\, dx &= \int_0^\infty \left(\frac{y}{\beta}\right)^{\alpha-1} e^{-y}\frac{dy}{\beta} \\ &= \frac{1}{\beta^\alpha}\int_0^\infty y^{\alpha-1}e^{-y}dy \\ &= \frac{\Gamma(\alpha)}{\beta^\alpha}\end{aligned}$$

Distribuição Gama

Definição 2 (Distribuição Gama): *Sejam α e β números positivos, então uma variável X segue a distruibuição gama com parâmetros α e β se X tiver uma f.d.p. contínua dada por:*

$$f(x;\alpha,\beta) = \begin{cases}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\exp(-\beta x), & \textit{para } x > 0 \\ 0, & \textit{caso contrário}\end{cases}$$



A prova da normalização de $\Gamma(\alpha)$ sai diretamente da equação (2)

Teorema 4: *Temos que:*

$$E[X] = \frac{\alpha}{\beta}$$

e

$$Var(X) = \frac{\alpha}{\beta^2}$$

Prova: Temos que para $k = 1, 2, ...$

$$\begin{aligned}\mathsf{E}\left[X^k\right] &= \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)}x^k x^{\alpha-1}\exp(-\beta x)\, dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{k+\alpha-1}\exp(-\beta x)\, dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(k+\alpha)}{\beta^{k+\alpha}} \\ &= \frac{\Gamma(k+\alpha)}{\Gamma(\alpha)\,\beta^k}\end{aligned}$$

Assim:

$$\mathsf{E}[X] = \frac{\Gamma(1+\alpha)}{\Gamma(\alpha)\,\beta} = \frac{(\alpha)\Gamma(\alpha)}{\Gamma(\alpha)\,\beta} = \frac{\alpha}{\beta}$$

e, por outro lado:

$$\mathsf{E}\left[X^2\right] = \frac{\Gamma(2+\alpha)}{\Gamma(\alpha)\,\beta^2} = \frac{(1+\alpha)(\alpha)\Gamma(\alpha)}{\Gamma(\alpha)\,\beta^2} = \frac{(1+\alpha)\alpha}{\beta^2}$$

A variância é dada por:

$$\begin{aligned}\mathsf{Var}(X) = \mathsf{E}\left[X^2\right] - \mathsf{E}[X] &= \frac{(1+\alpha)\alpha}{\beta^2} - \frac{\alpha^2}{\beta^2} \\ &= \frac{\alpha}{\beta^2}\end{aligned}$$

Teorema 5: *Se as variáveis X_1, X_2, ..., X_k forem independentes e se cada X_i $(i = 1, 2, ..., k)$ seguirem uma distribuição Gama com parâmetros α_i e β, então a soma $Y = X_1 + X_2 + ... + X_k$ segue uma distribuição Gama com parâmetros $\alpha = \alpha_1 + \alpha_2 + ... + \alpha_k$ e β.*

Não será dada aqui uma prova formal deste teorema. Entretanto, podemos mostrar que o valor esperado e a variância de Y são compatíveis, de fato, com o valor esperado e a variância de uma distribuição Gama com os parâmetros descritos acima. O valor esperado é dado por:

$$\begin{aligned}\mathsf{E}[Y] &= \mathsf{E}[X_1] + \mathsf{E}[X_2] + ...\mathsf{E}[X_k] \\ &= \frac{\alpha_1}{\beta} + \frac{\alpha_2}{\beta} + ... + \frac{\alpha_k}{\beta} \\ &= \frac{\alpha_1 + \alpha_2 + ...\alpha_k}{\beta},\end{aligned}$$

que é o valor esperado de uma distribução Gama com $\alpha = \alpha_1 + \alpha_2 + ... + \alpha_k$ e β. Por outro lado:

$$\begin{aligned}\mathsf{Var}(Y) &= \mathsf{Var}(X_1) + \mathsf{Var}(X_2) + ...\mathsf{Var}(X_k) \\ &= \frac{\alpha_1}{\beta^2} + \frac{\alpha_2}{\beta^2} + ... + \frac{\alpha_k}{\beta^2} \\ &= \frac{\alpha_1 + \alpha_2 + ...\alpha_k}{\beta^2},\end{aligned}$$

Distribuição Exponencial

Este é um caso particular da Distribuição Gama, em que $\alpha = 1$.

Definição 3 (Distribuição Exponencial): *Seja $\beta > 0$, uma variável aleatória X segue uma distribuição exponencial com parâmetro β se X seguir a f.d.p. contínua dada por:*

$$f(x;\beta) = \begin{cases}\beta\exp(-\beta x), & \textit{para } x > 0 \\ 0, & \textit{caso contrário}\end{cases}$$

Podemoves ver da definição da função Gama que E $[X] = \frac{1}{\beta}$ e Var $(X) = \frac{1}{\beta^2}$.

Distribuição Exponencial não tem Memória

Teorema 6: *Seja X uma variável aleatória que segue uma f.d.p. exponencial com parâmetro $\beta > 0$. Seja $t > 0$, então para $h > 0$ temos:*

$$Pr(X \geq t + h | X \geq t) = Pr(X \geq h)$$

O teorema acima está a nos dizer que a probabilidade da variável X ser maior do que um valor $t + h$ é exatamente a mesma da variável ser maior do que h, não importa o valor de t.

Prova: Temos:

$$\begin{aligned}
\Pr(X \geq t) &= \int_t^\infty \beta \exp(-\beta x) dx \\
&= \beta - \frac{\exp(-\beta x)}{\beta} \bigg|_t^\infty \\
&= \exp(-\beta t)
\end{aligned}$$

Se lembrarmos da f.d.p. condicional:

$$\Pr(B|A) = \frac{\Pr(A \cup B)}{\Pr(B)}$$

temos:

$$\begin{aligned}
\Pr(X \geq t + h | X \geq t) &= \frac{\Pr([X \geq t + h] \cup [X \geq t])}{\Pr(X \geq t)} \\
&= \frac{\Pr(X \geq t + h)}{\Pr(X \geq t)} \\
&= \frac{\exp(-\beta(t + h))}{\exp(-\beta t)} \\
&= \exp(-\beta h) \\
&= \Pr(X \geq h)
\end{aligned}$$

Exemplo: Considere uma empresa que quer saber informações sobre o tempo de vida de suas lâmpadas. Ela observa o tempo de vida $X_i = 1, 2, ..., n$ de n lâmpadas. Suponha que o tempo de vida de uma lâmpada seja dado por uma distribuição exponencial, com parâmetro β. O que se pode dizer sobre a variável Y_1, que é o tempo necessário para que a primeira lâmpada falhe? E sobre os tempos Y_2, Y_3 e assim por diante?

A distribuição de Y_1 é o valor mínimo das X_i variáveis aleatórias registradas para cada lâmpada. Assim fica fácil de se descobrir a distribuição de Y_1.

Teorema 7: *Se X_1, X_2, .., X_n forem n variáveis aleatórias que seguem uma distribuição exponencial com parâmetro β, então a distribuição de $Y_1 = \min\{X_1, X_2, ..., X_n\}$ terá a distribuição exponencial com parâmetro $n\beta$.*

Prova: Para cada número $t > 0$:

$$\begin{aligned}
\Pr(Y_1 > t) &= \Pr(X_1 > t, X_2 > t, ..., X_n > t) \\
&= \Pr(X_1 > t) \Pr(X_2 > t) \ldots \Pr(X_n > t) \\
&= e^{-\beta t} e^{-\beta t} \ldots e^{-\beta t} \\
&= e^{-(n\beta)t}
\end{aligned}$$

Podemos ver que este resultado bate com aquele visto por uma distribuição exponencial com parâmetro $n\beta$.

Como o processo é sem memória, então temos que a após a primeira lâmpada queimar, é como se o processo iniciasse novamente em $t = 0$. Assim, $Y_2 - Y_1$ terá uma distribuição expoencial com parâmetro $(n - 1)\beta$, com as $n - 1$ lâmpadas restantes. O raciocínio é o mesmo para os outros valores de Y_i.

Teorema 8: *Se X_1, X_2, .., X_n forem n variáveis aleatórias que seguem uma distribuição exponencial com parâmetro β, e se as variáveis $Z_1, Z_2, ..., Z_n$ forem as variáveis X_i arranjadas em ordem de acontecimento, então a distribuição de $Y_k = Z_k - Z_{k-1}$, com $k = 2, 3, ..., n$ terá a distribuição exponencial com parâmetro $(n + 1 - k)\beta$.*

A prova segue do teorema anterior.

Relação com a Distribuição de Poisson

Consideremos um exemplo onde uma fonte radioativa emite partículas que são detectadas. Estas partículas seguem uma distribuição do tipo Poisson com parâmetro β. Seja Z_k o tempo que demora para que a k-ésima partícula detectada demora deste o início da contagem, qual é a distribuição de Z_1? Qual é a distribuição de $Y_k = Z_k - Z_{k-1}$, para $k \geq 2$?

Primeiramente devemos notar que há uma diferença crucial entre este problema e o problema das lâmpadas sugerido anteriormente. Lá o número de lâmpadas era fixo. Aqui não sabemos a princípio qual será o número n de contagens. De fato, o número de contagens dependerá, entre outras coisas, do tempo de observação.

Teorema 9: *Se o número de acontecimentos de determinado experimento segue uma distribuição de Poisson com parâmetro β, então se Z_k é o tempo entre o início do experimento e o tempo em que o k-ésimo evento ocorre, e definirmos $Y_k = Z_k - Z_{k-1}$ então Y_1, Y_2, ..., seguem uma distribuição exponencial com parâmetro β.*

Prova: Se $t > 0$ e X é o número de acontecimentos até o tempo t, então se $Y_1 \leq t$ é porque ao menos um acontecimento já ocorreu até o tempo t. Como sabemos que X segue uma distribuição de Poisson com parâmetro β Assim:

$$\Pr(Y_1 \leq t) = \Pr(X \geq 1) = 1 - \Pr(X = 0) = 1 - e^{-\beta t}$$

Como vimos anteriormente, para a distribuição exponencial $\Pr(Y \geq t) = e^{-\beta t}$. Assim podemos ver que Y_1 segue uma distribuição exponencial com parâmetro β. Por causa da falta de memória desta distribuição, podemos ver que o mesmo deve acontecer para Y_2, Y_3, ...

Distribuições de χ^2

Distribuições do tipo χ^2 são distribuições Gama com parâmetros $\alpha = m/2$ e $\beta = 1/2$. Diz-se que m é o chamados *número de graus de liberdade*. Assim, temos:

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{2^{m/2}\,\Gamma(m/2)}\, x^{(m/2)-1} \exp(-x/2), & \text{para } x > 0 \\ 0, & \text{caso contrário} \end{cases}$$

Podemos ver facilmente que $\mathrm{E}[X] = m$ e $\mathrm{Var}(X) = 2m$.

Teorema 10: *0 Sejam $X_1, ..., X_k$, k variáveis independentes e que X_i siga uma distribuição de χ^2 com m_i graus de liberdade, então $X_1 + ... + X_k$ tem a distribuição de χ^2 com $m_1 + ... + m_k$ graus de liberade.*

Este teorema segue diretamente da propriedade vinda da distribuição Gama.

Teorema 11: *Seja X uma variável que tenha distribuição normal padrão. Então a variável aleatória $Y = X^2$ tem distribuição χ^2 com um grau de liberdade.*

Prova: Seja $f(y)$ e $F(y)$ a f.d.p e a c.d.f. de Y. Como X segue uma distribuição normal padrão, temos que $\phi(x)$ e $\Phi(X)$ são, respectivamente, a f.d.p. e a c.d.f de X. Assim, para $y > 0$:

$$\begin{aligned} F(y) &= \Pr(Y \leq y) \\ &= \Pr(X^2 \leq y) \\ &= \Pr(-y^{1/2} \leq X \leq y^{1/2}) \\ &= \Phi(y^{1/2}) - \Phi(-y^{1/2}) \end{aligned}$$

Agora, vale lembrar que $f(y) = F'(y)$, assim, derivando ambos os lados da equação de cima e aplicando a regra da cadeia, temos:

$$f(y) = \phi(y^{1/2})\left[\frac{1}{2}y^{-1/2}\right] - \phi(-y^{1/2})\left[-\frac{1}{2}y^{-1/2}\right]$$

Como $\phi(y^{1/2}) = \phi(-y^{1/2}) = (2\pi)^{-1/2}\, e^{-y/2}$, temos:

$$f(y) = \frac{1}{\sqrt{2\pi}}\, y^{-1/2}\, e^{-y/2}, \quad \text{para } y > 0.$$

Podemos ver que esta é uma f.d.p. de uma distribuição χ^2 com um grau de liberdade.

Corolário 1: *Se as variáveis aleatórias X_1, X_2, ..., X_m são independentes e identicamente distribuídas com distribuição normal, então a soma dos quadrados $X_1^2 + X_2^2 + ... + X_m^2$ tem distribuição χ^2 com m graus de liberdade.*

Distribuição da Média de Amostra e da Variância de Amostra

Suponha que tenhamos uma série de n amostras: X_1, X_2, ..., X_n vindos de uma distribuição normal. Estatísticas relevantes são dadas pela média da amostra $\bar{X}_n$ e pela variância da amostra S^2. O valor mais provável de μ é X_n e o valor mais provável da variância é $(1/n)\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$.

Sabemos de antes que a média $\bar{X}_n$ possui uma distribuição normal com média μ e com variância σ^2/n. Por outro lado vimos também que se um conjunto de variáveis Y_i for extraído de uma distribuição normal, então a variável $\sum_{i=1}^{n} Y_i^2$ seguirá uma distribuição do tipo χ^2 com n graus de liberdade. Se fizermos $Y_i = (X_i - \mu)/\sigma$, então temos que $\sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2$ seguirá uma distribuição do tipo χ^2 com n graus de liberdade.

Teorema 12: *Suponha que X_1, X_2, .., X_n seja uma amostra aleatória de uma distribuição normal com média μ e com variância σ. Então a média da amostra $\bar{X}_n$ e a variância da amostra $(1/n)\sum_{i=1}^{n}(X_i - \bar{X}_n)$ são variáveis independentes, $\bar{X}_n$ tem distribuição normal com média μ e variância σ^2/n e $\sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2$ tem a distribuição χ^2 com $n - 1$ graus de liberdade.*