

$$D^2(B) = \frac{1}{100}[(0,7-1)^2 \cdot 10 + (0,8-1)^2 \cdot 10 + (0,9-1)^2 \cdot 21 \\ + (1,0-1)^2 \cdot 18 + (1,1-1)^2 \cdot 15 + (1,2-1)^2 \cdot 22 \\ + (1,3-1)^2 \cdot 4] = 0,0290.$$

Ponieważ $D^2(A) < D^2(B)$, więc należy wybrać partię A .

Przykład 3.14. Obliczyć wariancję rozkładu jednostajnego.

Ponieważ

$$D^2(X) = E(X^2) - (E(X))^2 = m_2 - m_1^2,$$

a na podstawie zadania 6 po p. 3.6 wiemy, że

$$m_1 = E(X) = \frac{a+b}{2},$$

więc wystarczy obliczyć m_2 . Mamy

$$m_2 = \int_{-\infty}^{+\infty} x^2 f(x) = \frac{1}{b-a} \int_a^b x^2 dx = \frac{1}{3} \cdot \frac{1}{b-a} \cdot x^3 \Big|_a^b = \frac{1}{3}(a^2 + ab + b^2).$$

Zatem

$$D^2(X) = \frac{1}{3}(a^2 + ab + b^2) - \frac{1}{4}(a^2 + 2ab + b^2) = \frac{(a-b)^2}{12}.$$

Definicja 3.28. Jeżeli X oznacza zmienną losową o odchyleniu standardowym $D(X) = \sigma$, to zmienną losową $Y = X/\sigma$ nazywamy *unormowaną (znormalizowaną)*.

Twierdzenie 3.14. *Wariancja zmiennej losowej unormowanej jest równa 1.*

Dowód. Na podstawie twierdzenia 3.12 mamy

$$D^2(Y) = D^2\left(\frac{X}{\sigma}\right) = \frac{1}{\sigma^2} D^2(X) = \frac{1}{\sigma^2} \sigma^2 = 1. \quad \blacksquare$$

Definicja 3.29. Jeśli X oznacza zmienną losową o wartości oczekiwanej $E(X) = m$ i odchyleniu standardowym $D(X) = \sigma$, to zmienną losową $Y = \frac{X-m}{\sigma}$ nazywamy *standaryzowaną*.

Twierdzenie 3.15. *Wartość oczekiwana zmiennej losowej standaryzowanej jest równa zero, a wariancja jest równa 1.*

Dowód. Na podstawie twierdzeń 3.5 i 3.7 mamy

$$E(Y) = E\left(\frac{X-m}{\sigma}\right) = \frac{E(X)-m}{\sigma} = \frac{m-m}{\sigma} = 0,$$

a na podstawie twierdzeń 3.11 i 3.12 otrzymujemy

$$D^2(Y) = D^2\left(\frac{X-m}{\sigma}\right) = D^2\left(\frac{X}{\sigma} - \frac{m}{\sigma}\right) = \frac{1}{\sigma^2} D^2(X) = \frac{\sigma^2}{\sigma^2} = 1. \quad \blacksquare$$

Odchylenie standardowe nie jest jedynym parametrem charakteryzującym rozrzut. Jest jeszcze tzw. odchylenie przeciętne.

Definicja 3.30. *Odchyleniem przeciętnym $\alpha = d(X)$ zmiennej losowej X nazywamy wyrażenie*

$$\alpha = d(X) = \sum_k |x_k - E(X)| p_k \quad \text{dla zmiennej losowej skokowej,}$$

$$\alpha = d(X) = \int_{-\infty}^{+\infty} |x - E(X)| f(x) dx \quad \text{dla zmiennej losowej ciągłej,}$$

tj.

$$\alpha = E(|X - E(X)|).$$

Twierdzenie 3.16. *Dla dowolnej zmiennej losowej X zachodzi*

$$d(X) \leq D(X).$$

Dowód. Z nierówności Schwarza dla $Y = 1$ mamy

$$d^2(X) = (E(|X - E(X)|))^2 \leq E((X - E(X))^2) = D^2(X). \quad \blacksquare$$

Zadania

1. Rzucamy sześcienną kostką do gry. Obliczyć wariancję liczby wyrzuconych oczek na kostce.
2. Rzucamy dwiema sześciennymi kostkami do gry. Obliczyć wariancję sumy liczb oczek otrzymanych na kostkach.
3. Obliczyć wariancję rozkładu Bernoulliego.
4. Obliczyć wariancję rozkładu Poissona.
5. Rzucamy dwie kości do gry. Oznaczmy przez X_1 zmienną losową przyjmującą wartości równe liczbie oczek na pierwszej kostce, a przez X_2 zmienną losową przyjmującą wartość 1, o ile na pierwszej kostce jest piątka i na drugiej kostce jest piątka, natomiast wartość 0 w pozostałych przypadkach.
 - A. Znaleźć rozkład zmiennej losowej $X = X_1 + X_2$.
 - B. Wyznaczyć dystrybuantę zmiennej losowej X i wykonać jej wykres.
 - C. Obliczyć $P(4 \leq X < 6)$ oraz zaznaczyć obliczone prawdopodobieństwo na wykresie dystrybuanty.
 - D. Obliczyć wartość oczekiwaną i odchylenie standardowe zmiennej losowej X .

6. Myśliwy ma trzy naboje i strzela do chwili trafienia do celu lub do chwili wystrzelenia wszystkich naboji. Liczba naboji jest zmienną losową.
- A. Podać rozkład tej zmiennej wiedząc, że prawdopodobieństwo trafienia celu przy każdym strzale jest równe 0,8.
- B. Obliczyć wartość oczekiwaną i odchylenie standardowe.
7. A. Dla jakiej wartości C funkcja

$$f(x) = \begin{cases} Cx^2 & \text{dla } 0 \leq x \leq 3, \\ 0 & \text{dla } x < 0 \text{ i } x > 3 \end{cases}$$

jest gęstością prawdopodobieństwa?

- B. Znaleźć dystrybucję wyznaczonego rozkładu i wykonać jej wykres.
- C. Obliczyć $P(1 \leq X < 4)$ i zaznaczyć wartość obliczonego prawdopodobieństwa na wykresie dystrybucji.
- D. Obliczyć wartość oczekiwaną i odchylenie standardowe.
8. Zmienna losowa X przyjmuje wartości $-2, 2$ i 3 z prawdopodobieństwami odpowiednio $0,25, 0,5$ i $0,25$.
- A. Znaleźć rozkład zmiennej losowej $Y = X^2$.
- B. Obliczyć $E(Y)$ i $D(Y)$.

3.8. Wskaźniki zmienności, skośności i skupienia

Definicja 3.31. *Wskaźnikiem (współczynnikiem) zmienności zmiennej losowej X nazywamy wielkość*

$$V = \frac{D(X)}{E(X)},$$

przy czym $E(X) \neq 0$.

Definicja 3.32. Mówimy, że zmienna losowa ma rozkład prawdopodobieństwa *symetryczny* względem prostej $x = a$, jeśli

- a) w przypadku zmiennej losowej skokowej o punktach skokowych $\{x_k\}$ dla każdego punktu $x_i \leq a$ istnieje taki punkt $x_j \geq a$, że

$$P(X = x_i) = P(X = x_j) \text{ oraz } a - x_i = x_j - a,$$

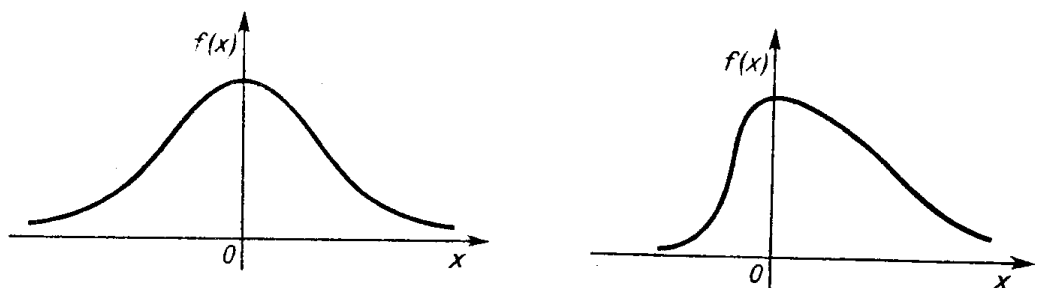
- b) w przypadku zmiennej losowej ciągłej o gęstości $f(x)$ zachodzi

$$f(a - x) = f(a + x)$$

dla każdej wartości x w punktach ciągłości funkcji $f(x)$.

Prosta o równaniu $x = a$ nazywa się *osią symetrii*.

Rozkład symetryczny i niesymetryczny przedstawiono na następujących rysunkach:



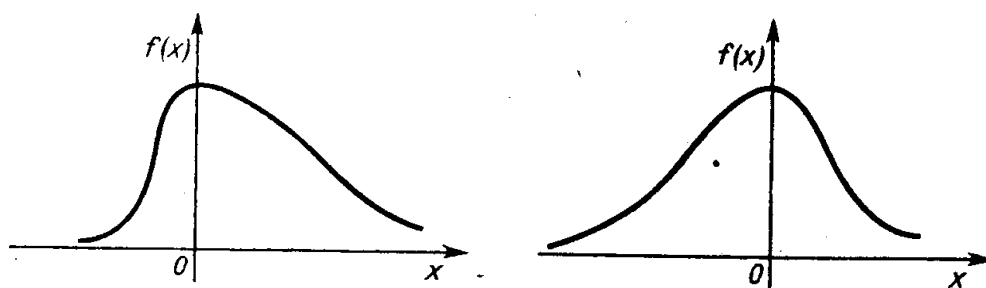
Jeśli rozkład prawdopodobieństwa jest symetryczny, to momenty centralne rzędu nieparzystego, o ile istnieją, są równe zero. Z tego powodu za miarę asymetrii przyjęto wskaźnik zdefiniowany następująco:

Definicja 3.33. Wskaźnikiem asymetrii (współczynnikiem skośności) nazywamy wyrażenie

$$\gamma_1 = \frac{\mu_3}{D^3(X)},$$

gdzie $D(X) > 0$.

Gdy $\gamma_1 > 0$, to mówimy, że asymetria rozkładu prawdopodobieństwa jest dodatnia (zob. rysunek z lewej poniżej), a gdy $\gamma_1 < 0$ – ujemna (zob. rysunek z prawej poniżej). Czasami jako miara skośności używany jest też wskaźnik $\beta_1 = \gamma_1^2$.



Definicja 3.34. Wskaźnikiem skupienia (współczynnikiem spłaszczenia lub kurtozą) nazywamy wyrażenie

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{D^4(X)},$$

a ekscesem – wyrażenie

$$\gamma_2 = \beta_2 - 3.$$

Eksces rozkładu normalnego (będzie dalej) jest równy zero i dlatego jest on wskaźnikiem umożliwiającym porównanie skupienia badanego rozkładu ze względu na rozkład normalny. Eksces dodatni wskazuje, że wykres badanego rozkładu jest wyższy i bardziej smukły (bardziej skupiony ze względu na oś poprowadzoną przez punkt $x_0 = E(X)$). Eksces ujemny ma znaczenie przeciwne.

3.9. Mediana i moda

Wskaźniki te, obok wartości oczekiwanej, charakteryzują położenie zbioru wartości zmiennej losowej.

Definicja 3.35. *Medianą Me zmiennej losowej X nazywamy liczbę x spełniającą związek*

$$P(X \leq x) \geq \frac{1}{2} \text{ i } P(X \geq x) \geq \frac{1}{2}.$$

Dla zmiennej losowej ciągłej o gęstości $f(x)$ i dystrybucji $F(x)$ nierówności te sprowadzają się do równania

$$F(x) = \frac{1}{2},$$

tj.

$$\int_{-\infty}^{x=Me} f(x) = \frac{1}{2}.$$

Definicja 3.36. *Modą Mo lub dominatą zmiennej losowej X nazywamy:*

- dla zmiennej losowej skokowej – wartość zmiennej losowej, której odpowiada największe prawdopodobieństwo,
- dla zmiennej losowej ciągłej – wartość, dla której gęstość przyjmuje maksimum lokalne.

Przykład 3.15. Wyznaczyć medianę i modę dla zmiennej losowej skokowej o rozkładzie podanym w poniższej tabeli.

x_k	2	5	7	10
$P(X = x_k)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{5}{9}$	$\frac{1}{9}$

Mediana $Me = 7$, bo liczba ta spełnia obie nierówności

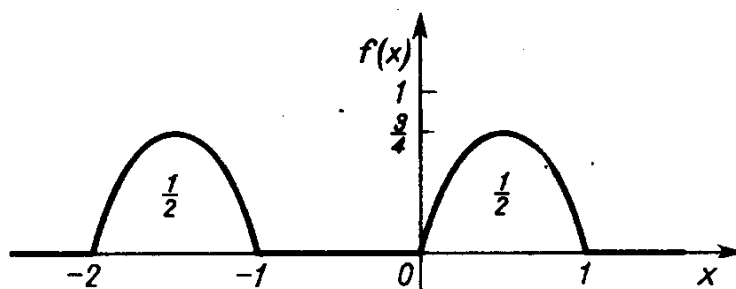
$$P(X \leq 7) = \frac{8}{9} > \frac{1}{2} \text{ i } P(X \geq 7) = \frac{6}{9} > \frac{1}{2}.$$

Bezpośrednio z tabeli mamy $Mo = 7$.

Przykład 3.16. Wyznaczyć medianę i modę zmiennej losowej o rozkładzie

$$f(x) = \begin{cases} 0 & \text{dla } -\infty < x \leq -2, \\ -3x^2 - 9x - 6 & \text{dla } -2 < x \leq -1, \\ 0 & \text{dla } -1 < x \leq 0, \\ -3x^2 + 3x & \text{dla } 0 < x \leq 1, \\ 0 & \text{dla } 1 < x < +\infty. \end{cases}$$

Szkic wykresu funkcji $f(x)$ jest następujący:



Medianą jest każda liczba x z przedziału domkniętego $[-1, 0]$, gdyż dla tych wartości mamy

$$\int_{-\infty}^x f(t) dt = \int_{-2}^{-1} (-3t^2 - 9t - 6) dt = \frac{1}{2}.$$

Funkcja gęstości ma ekstrema lokalne w punktach $-\frac{3}{2}$ i $\frac{1}{2}$. Zatem

$$Mo = -\frac{3}{2} \text{ i } Mo = \frac{1}{2}.$$

Zadania

1. Wyznaczyć medianę, modę oraz wskaźniki asymetrii, skupienia i eksces w przypadku, gdy zmienna losowa skokowa ma rozkład podany w poniższej tabeli

x_k	-3	-2	-1	0	1	2	3
$P(X=x_k)$	0,05	0,12	0,18	0,32	0,23	0,07	0,03

2. Wyznaczyć medianę zmiennej losowej o rozkładzie wykładniczym.
3. Rzucamy dwie kości do gry. Oznaczmy przez X_1 zmienną losową przyjmującą wartości równe liczbie oczek na pierwszej kostce, a przez X_2 zmienną losową przyjmującą wartość 1, o ile na pierwszej kostce jest piątka i na drugiej kostce jest piątka, natomiast wartość 0 w pozostałych przypadkach. Wyznaczyć medianę i modę zmiennej losowej $X = X_1 + X_2$.

4. Gęstość prawdopodobieństwa zmiennej losowej X dana jest wzorem

$$f(x) = \begin{cases} \frac{1}{9}x^2 & \text{dla } 0 \leq x \leq 3, \\ 0 & \text{dla } x < 0 \text{ i } x > 3 \end{cases}$$

Obliczyć Me i Mo .

5. Zmienna losowa X przyjmuje wartości $-2, 2$ i 3 z prawdopodobieństwami odpowiednio $0,25, 0,5$ i $0,25$. Obliczyć medianę i modę zmiennej losowej $Y = X^2$.

3.10. Rozkład normalny

Definicja 3.37. Zmienną losową o gęstości

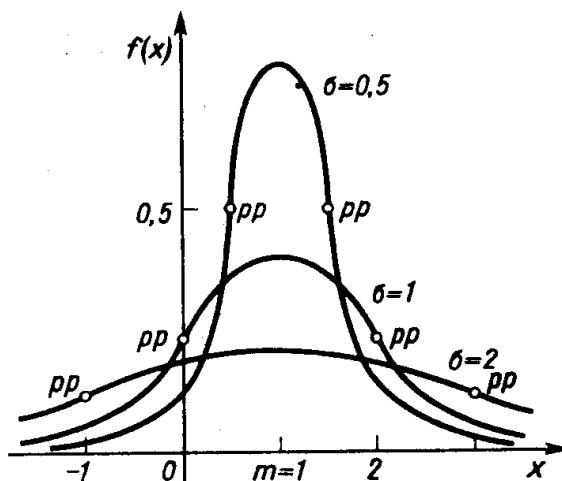
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right], \quad -\infty < x < +\infty, \quad \sigma > 0,$$

gdzie m oznacza wartość oczekiwaną, a σ – odchylenie standardowe, nazywamy zmienną losową o *rozkładzie normalnym* lub *rozkładzie Moivre'a-Gaussa*. Rozkład ten oznaczamy $N(m, \sigma)$.

Badając przebieg zmienności funkcji $f(x)$ (danej powyższym wzorem) można stwierdzić, że:

- a) funkcja jest rosnąca dla $x < m$ i malejąca dla $x > m$,
- b) w punkcie $x = m$ funkcja osiąga maksimum równe $\frac{1}{\sqrt{2\pi}\sigma}$,
- c) w punktach o odciętych $x_1 = m - \sigma$ i $x_2 = m + \sigma$ funkcja ma punkty przegięcia, których rzędne są równe $\frac{1}{\sqrt{2\pi e}\sigma}$.

Ponadto okazuje się, że wartość m jest modą i medianą rozkładu. Wykresy funkcji $f(x)$ dla $m = 1$ i różnych wartości σ są przedstawione na poniższym rysunku.



W szczególnym przypadku, gdy $m = 0$ i $\sigma = 1$ mówimy, że zmienna losowa ma *standaryzowany rozkład normalny*. Wówczas

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{x^2}{2}\right), \quad -\infty < x < +\infty.$$

Dystrybuantę standaryzowanego rozkładu normalnego nazywamy *funkcją* lub *całką Laplace'a* i oznaczamy symbolem $\Phi(x)$. Ma ona następującą postać:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt.$$

Nie jest to całka elementarna i jej obliczenie jest żmudne. Dlatego korzysta się z tablic wartości funkcji $\Phi(x)$.

Obliczenie wartości dystrybuanty $F(x)$ rozkładu $N(m, \sigma)$ przeprowadza się stosując podstawienie $\frac{t-m}{\sigma} = u$. Wówczas

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m}{\sigma}} \exp\left(-\frac{u^2}{2}\right) du = \Phi\left(\frac{x-m}{\sigma}\right).$$

Jeśli oznaczymy $\frac{x-m}{\sigma} = t$, to mamy $F(x) = \Phi(t)$.

Z symetrii gęstości standaryzowanego rozkładu normalnego wynika, że

$$\Phi(-x) = 1 - \Phi(x). \quad (*)$$

Mamy bowiem

$$P(X < -x) + P(X = -x) + P(X > -x) = 1.$$

Ale $P(X = -x) = 0$, bo rozkład jest ciągły. Zatem

$$P(X < -x) = 1 - P(X > -x),$$

a ponieważ gęstość jest symetryczna, co oznacza, że

$$P(X > -x) = P(X < x),$$

a więc

$$P(X < -x) = 1 - P(X < x),$$

co kończy dowód równości, bo $P(X < -x) = \Phi(-x)$ i $P(X < x) = \Phi(x)$.

Przykład 3.17. Zmienna losowa X podlega rozkładowi $N(5, 2)$. Obliczyć $P(X < 3,6)$.

Szukane prawdopodobieństwo jest równe polu pod krzywą gęstości w przedziale $(-\infty; 3,6)$. Nierówność $X < 3,6$ jest równoważna nierówności

$$\frac{X-5}{2} < \frac{3,6-5}{2}.$$

Zatem

$$\begin{aligned}
 P(X < 3,6) &= P\left(\frac{X-5}{2} < \frac{3,6-5}{2}\right) = P(T < -0,7) \\
 &= \Phi(-0,7) = 1 - \Phi(0,7) \approx 1 - 0,7580 = 0,2420,
 \end{aligned}$$

przy czym wartość $\Phi(0,7)$ została odczytana z tablic.

Przykład 3.18. Zmienna losowa X podlega rozkładowi $N(1, 2)$. Obliczyć $P(|X| < 2,4)$. Wyznać stałą a tak, by $P(|X-1| < a) \approx 0,95$.

Po standaryzacji zmiennej losowej i uwzględnieniu wzoru (*) mamy

$$\begin{aligned}
 P(|X| < 2,4) &= P(-2,4 < X < 2,4) = P\left(\frac{-2,4-1}{2} < \frac{X-1}{2} < \frac{2,4-1}{2}\right) \\
 &= P(-1,7 < T < 0,7) = \Phi(0,7) - \Phi(-1,7) = \Phi(0,7) + \Phi(1,7) - 1 \approx 0,7134.
 \end{aligned}$$

Aby znaleźć stałą a , zauważmy, że

$$P(|X-1| < a) = P\left(-\frac{a}{2} < \frac{X-1}{2} < \frac{a}{2}\right) = 2\Phi\left(\frac{a}{2}\right) - 1.$$

Stąd otrzymujemy równanie

$$2\Phi\left(\frac{a}{2}\right) - 1 = 0,95,$$

czyli $\Phi\left(\frac{a}{2}\right) = 0,975$. Korzystając z tablic mamy $a \approx 3,9$.

Zadania

1. Zmienna losowa podlega rozkładowi $N(2, 4)$. Obliczyć $P(|X| > 6)$.
2. Zmienna losowa podlega rozkładowi $N(3, 5)$. Obliczyć $P(|X-1| \geq 1)$.
3. Zmienna losowa X podlega rozkładowi $N(4, 9)$. Obliczyć:
 - a) $P(X > 13)$,
 - b) $P(|X-2| < 14)$,
 - c) $P(|X-3| > 11)$.
4. Zmienna losowa podlega rozkładowi $N(2, 4)$.
 - A. Wyznaczyć wartość a tak, aby $P(|X-2| < a) = 0,9$.
 - B. Wyznaczyć wartość k tak, aby $P(|X-2| < 4k) = 0,8$.

3.11. Rozkład logarytmiczno-normalny

Definicja 3.38. Rozkład zmiennej losowej X , której logarytm naturalny ($\ln X$) ma rozkład normalny, nazywamy *rozkładem logarytmiczno-naturalnym*.

Gęstość rozkładu logarytmiczno-normalnego można otrzymać z gęstości rozkładu normalnego zmiennej losowej $Y = \ln X$. Wówczas zmienna losowa $X = \exp Y$ ma opisywany rozkład, którego gęstość dana jest wzorem

$$f(x) = \begin{cases} 0 & \text{dla } x \leq 0, \\ \frac{1}{\sqrt{2\pi}\sigma} \cdot \frac{1}{x} \exp\left(-\frac{(\ln x - m)^2}{2\sigma^2}\right) & \text{dla } x > 0. \end{cases}$$

Rozkład logarytmiczno-normalny oznaczamy przez $\Lambda(m, \sigma)$.

Przez zbadanie przebiegu zmienności gęstości $f(x)$ rozkładu $\Lambda(m, \sigma)$ można wykazać, że:

a) funkcja posiada maksimum w punkcie $x_0 = \exp(m - \sigma^2)$, które jest równe

$$\frac{1}{\sqrt{2\pi}\sigma \cdot \exp\left(m - \frac{\sigma^2}{2}\right)},$$

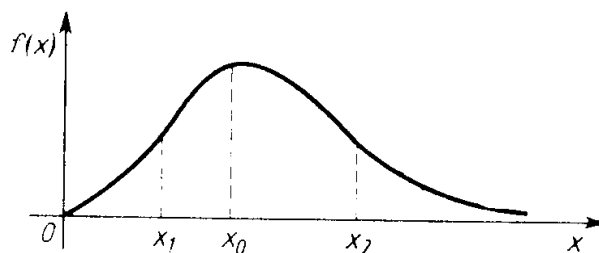
b) funkcja $f(x)$ jest rosnąca dla $0 < x < x_0$ i malejąca do zera dla $x > x_0$,

c) punkty

$$x_1 = \exp\left(m - \frac{2 + \sigma^2 + \sqrt{\Delta}}{2}\right) \text{ i } x_2 = \exp\left(m - \frac{2 + \sigma^2 - \sqrt{\Delta}}{2}\right),$$

gdzie $\Delta = (\sigma^2 - 2)^2 + 4$, są punktami przegięcia.

Ponadto okazuje się, że punkt x_0 jest także modą rozkładu. Wykres gęstości rozkładu logarytmiczno-normalnego jest przedstawiony na poniższym rysunku.



Przykład 3.19. Zmienna losowa X podlega rozkładowi $\Lambda(1, 2)$. Obliczyć $P(X < e^2)$.

Mamy

$$P(X < e^2) = \frac{1}{2\sqrt{2\pi}} \cdot \int_0^{e^2} x^{-1} \exp\left(-\frac{(\ln x - 1)^2}{2 \cdot 4}\right) dx = \frac{1}{2\sqrt{2\pi}} \int_0^{e^2} x^{-1} \exp\left[-\frac{1}{2} \cdot \left(\frac{\ln x - 1}{2}\right)^2\right] dx.$$

Podstawiając $\frac{\ln x - 1}{2} = u$ otrzymujemy

$$P(X < e^2) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\frac{1}{2}} \exp\left(-\frac{1}{2}u^2\right) du = \Phi(0,5) \approx 0,6915.$$

Przykład 3.20. Zmienna losowa X podlega rozkładowi $\Lambda(m, \sigma)$. Wyznaczyć momenty zwykłe rzędu l .

Zgodnie z definicją 3.23 mamy

$$m_l = \int_{-\infty}^{+\infty} x^l f(x) dx = \frac{1}{\sqrt{2\pi}\sigma} \cdot \int_0^{\infty} x^{l-1} \cdot \exp\left(-\frac{(\ln x - m)^2}{2\sigma^2}\right) dx.$$

Podstawiając $\frac{\ln x - m}{\sigma} = u$ otrzymujemy

$$\begin{aligned} m_l &= \frac{1}{\sqrt{2\pi}} \cdot \exp(lm) \cdot \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}(u^2 - 2l\sigma u)\right] du \\ &= \frac{1}{\sqrt{2\pi}} \cdot \exp(lm) \cdot \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}(u - l\sigma)^2 + \frac{1}{2}l^2\sigma^2\right] du \\ &= \frac{1}{\sqrt{2\pi}} \cdot \exp\left(lm + \frac{1}{2}l^2\sigma^2\right) \cdot \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}(u - l\sigma)^2\right] du. \end{aligned}$$

Stosując podstawienie $u - l\sigma = z$ mamy

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}(u - l\sigma)^2\right] du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{1}{2}z^2\right) dz = 1,$$

co wynika z własności funkcji gęstości dowolnego rozkładu normalnego (tutaj $N(0, 1)$). Zatem ostatecznie

$$m_l = \exp\left(lm + \frac{1}{2}l^2\sigma^2\right).$$

Zadania

1. Zmienna losowa X podlega rozkładowi $\Lambda(m, \sigma)$. Wyznaczyć medianę Me .
2. Wyznaczyć dystrybuantę rozkładu $\Lambda(m, \sigma)$.