

Monte Carlo RNG Algorithms

Bartosz Wędziński

11-12-2025

Spis treści

1 Symulacja dyskretnych zmiennych losowych	2
2 Symulacja ciągłych zmiennych losowych	4
3 Generowanie rozkładu normalnego metodą Boxa-Müllera	8
4 Metoda akceptacji	12

1 Symulacja dyskretnych zmiennych losowych

1.1 Funkcja realizująca generator liczb losowych o rozkładzie dwupunktowym: $\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$, gdzie $p \in [0, 1]$.

```
rbern <- function(n, p){
  if(p < 0 || p > 1) stop("p musi być w przedziale [0,1]")
  u <- runif(n)
  x <- ifelse(u < p, 1, 0)
  return(x)
}
```

1.2 Funkcja realizująca generator liczb losowych z rozkładu dwumianowego + histogram.

```
rbinom_custom <- function(N, n, p){
  x <- sapply(1:N, function(i) sum(rbern(n, p)))
  return(x)
}
```

Rozkład Bin(10, 0.3)

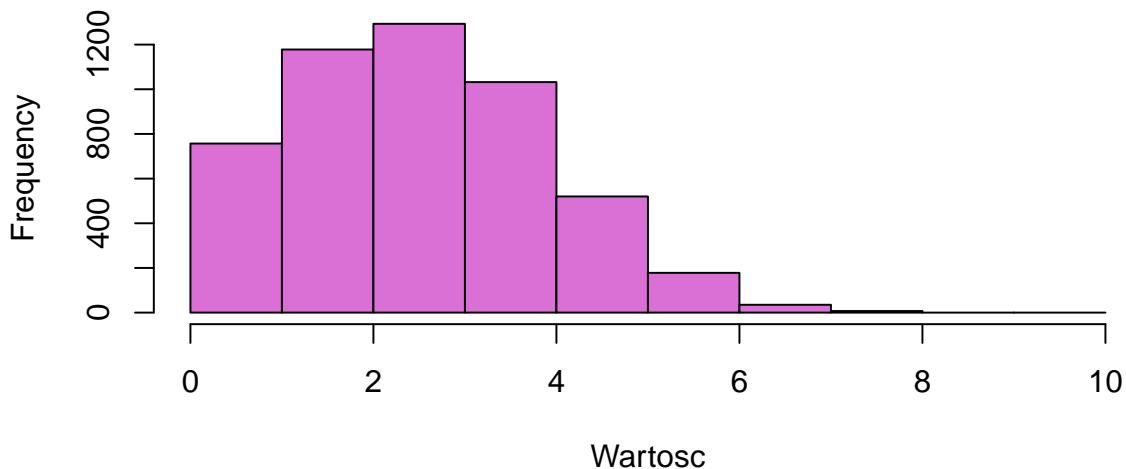


Figure 1: Histogram rozkładu dwumianowego Bin(10, 0.3).

Wyniki generowania zmiennej Bernoulliego zachowują się zgodnie z oczekiwaniami - przy dużej liczbie prób częstości zbliżają się do teoretycznych wartości.

1.3 Funkcja realizująca generator dla rozkładu Poissona + histogram.

```
rpois_custom <- function(n, lambda){
  if(lambda <= 0) stop("lambda musi być dodatnia")

  out <- numeric(n)
  for(j in 1:n){
    U <- runif(1)
    k <- 0
    p <- exp(-lambda)
    F <- p
    while(F < U){
      k <- k + 1
      p <- p * lambda / k
      F <- F + p
    }
    out[j] <- k
  }
  return(out)
}
```

Rozkład Poisson(4)

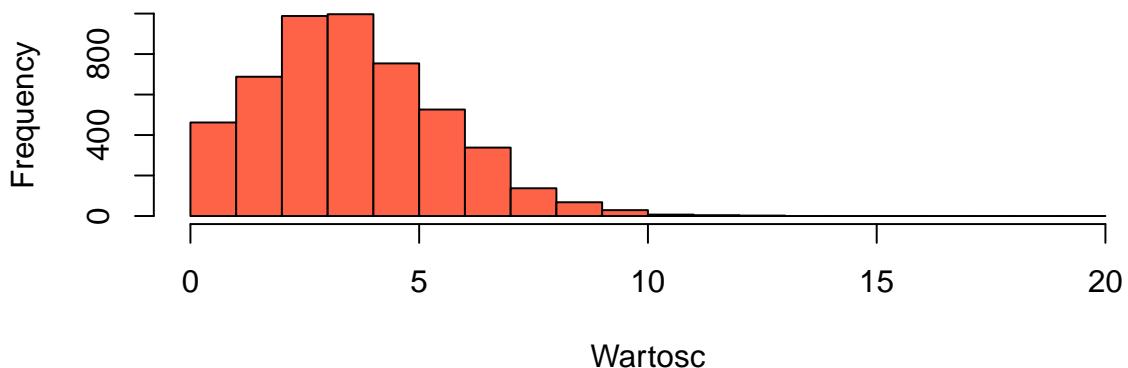


Figure 2: Histogram rozkładu Poissona (lambda=4).

Generator rozkładu Poissona również działa poprawnie: histogram ma typowy mocny pik w okolicach wartości równej parametrowi λ , a rozkład jest jednostronnie wydłużony w prawo. Różnice mieszczą się w granicach naturalnych odchyleń losowych.

2 Symulacja ciągłych zmiennych losowych

2.1 Generator rozkładu wykładniczego + histogram dla próby o rozmiarze $n = 2000$ wygenerowanej przy jego pomocy z krzywą gęstości rozkładu wykładniczego.

```
rexp_custom <- function(n, lambda){
  if(lambda <= 0) stop("lambda musi być dodatnia")
  u <- runif(n)
  x <- -log(1 - u) / lambda
  return(x)
}
```

Rozkład wykładniczy ($\lambda = 1$)

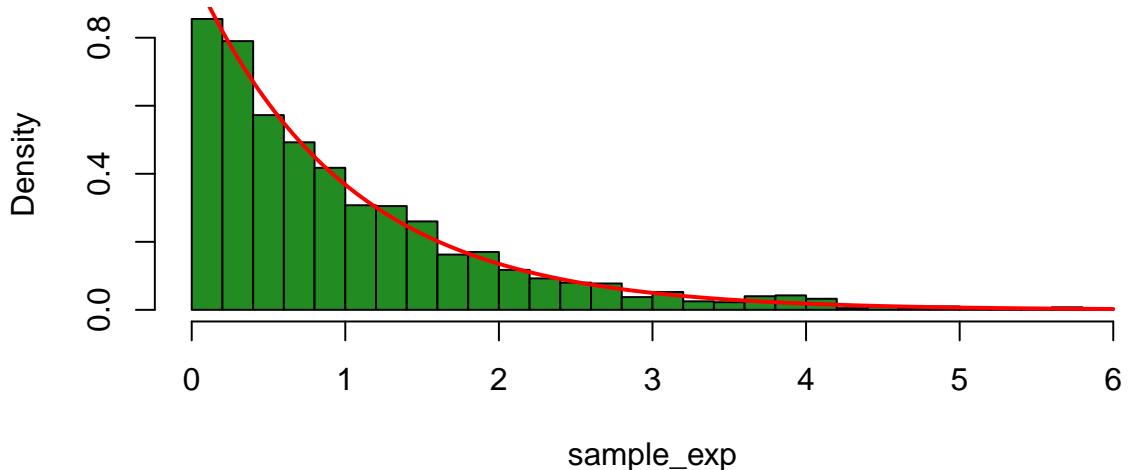


Figure 3: Rozkład wykładniczy - porównanie histogramu z gęstością.

Histogram generowanej zmiennej wykazuje klasyczny kształt malejący wykładniczo, zgodny z teoretyczną gęstością. Średnia empiryczna jest bliska wartości

$$\frac{1}{\lambda}$$

, co potwierdza poprawność generatora opartego na odwrotnej dystrybuancie.

2.2 Wyznaczanie dystrybuanty odwrotnej dla rozkładu Weibulla o funkcji gęstości

$$f(x) = \frac{k}{\lambda^k} x^{k-1} e^{-(x/\lambda)^k} \mathbf{1}_{(0,+\infty)}(x), \quad x \in \mathbb{R},$$

gdzie $k > 0$, $\lambda > 0$ + funkcja realizująca generator liczb losowych z rozkładu Weibulla + histogram dla próby o rozmiarze $n = 2000$ wygenerowanej przy jego pomocy z krzywą gęstości rozkładu Weibulla.

Wyprowadzenie dystrybuanty odwrotnej dla rozkładu Weibulla:

Dla rozkładu Weibulla o parametrach $k > 0$ oraz $\lambda > 0$ funkcja gęstości ma postać:

$$f(x) = \frac{k}{\lambda^k} x^{k-1} e^{-(x/\lambda)^k} \mathbf{1}_{(0,+\infty)}(x).$$

1. Dystrybuanta $F(x)$

Liczymy dystrybuantę jako całkę od zera do x :

$$F(x) = \int_0^x \frac{k}{\lambda^k} t^{k-1} e^{-(t/\lambda)^k} dt.$$

Wykonajmy podstawienie

$$u = \left(\frac{t}{\lambda}\right)^k \Rightarrow t = \lambda u^{1/k}, \quad dt = \lambda \frac{1}{k} u^{1/k-1} du.$$

Po podstawieniu:

$$\begin{aligned} F(x) &= \int_0^{(x/\lambda)^k} \frac{k}{\lambda^k} (\lambda u^{1/k})^{k-1} e^{-u} \lambda \frac{1}{k} u^{1/k-1} du \\ &= \int_0^{(x/\lambda)^k} e^{-u} du = \left[-e^{-u} \right]_0^{(x/\lambda)^k} = 1 - e^{-(x/\lambda)^k}. \end{aligned}$$

Zatem:

$$F(x) = 1 - e^{-(x/\lambda)^k}, \quad x > 0.$$

2. Odwrócenie dystrybuanty — $F^{-1}(u)$

Niech $u \in (0, 1)$. Mamy:

$$u = F(x) = 1 - e^{-(x/\lambda)^k}.$$

Przekształcamy krok po kroku:

$$\begin{aligned} e^{-(x/\lambda)^k} &= 1 - u, \\ -(x/\lambda)^k &= \ln(1 - u), \\ (x/\lambda)^k &= -\ln(1 - u), \\ x &= \lambda(-\ln(1 - u))^{1/k}. \end{aligned}$$

Ostatecznie:

$$F^{-1}(u) = \lambda(-\ln(1 - u))^{1/k}, \quad u \in (0, 1).$$

```
rweibull_custom <- function(n, k, lambda){
  if(k <= 0 || lambda <= 0) stop("k i lambda muszą być dodatnie")
  u <- runif(n)
  x <- lambda * (-log(1 - u))^(1/k)
  return(x)
}
```

Rozkład Weibulla (k=2, lambda=1)

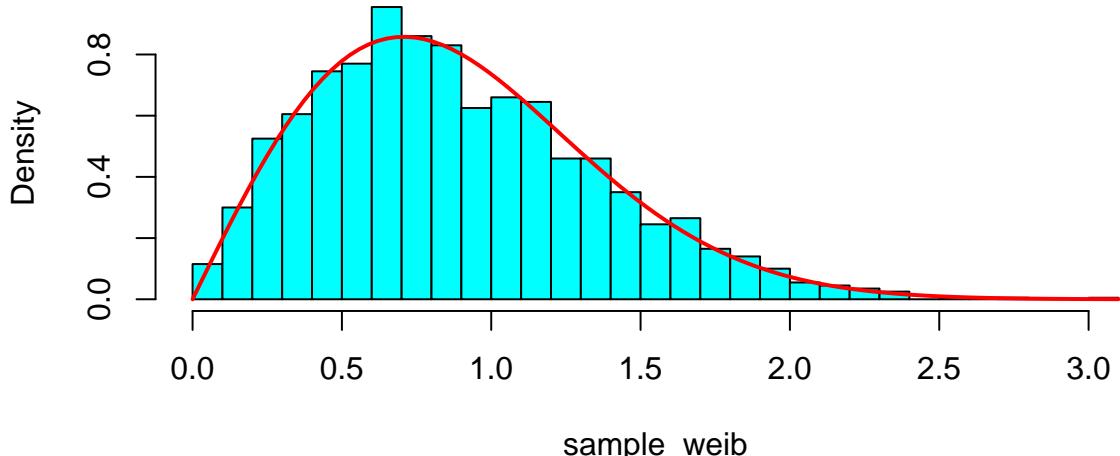


Figure 4: Rozkład Weibulla ($k=2$, $\lambda=1$).

Rozkład Weibulla ma wykres zależny od parametru kształtu k . Obserwujemy charakterystyczny wzrost gęstości, a następnie jej spadek. Wyniki liczbowo są stabilne i zbliżone do wartości teoretycznych.

2.3 Funkcja realizująca generator liczb losowych z rozkładu Laplace'a z parametrem $\lambda > 0$ o gęstości:

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x|}, \quad x \in \mathbb{R}.$$

+ histogram dla próby o rozmiarze $n = 2000$ wygenerowanej przy jego pomocy z krzywą gęstości rozkładu Laplace'a.

```
rlaplace_custom <- function(n, lambda){
  if(lambda <= 0) stop("lambda musi być dodatnia")
  u <- runif(n)
  x <- ifelse(u < 0.5,
               (1/lambda) * log(2*u),
               -(1/lambda) * log(2*(1-u)))
  return(x)
}
```

Rozkład Laplace'a (lambda=1)

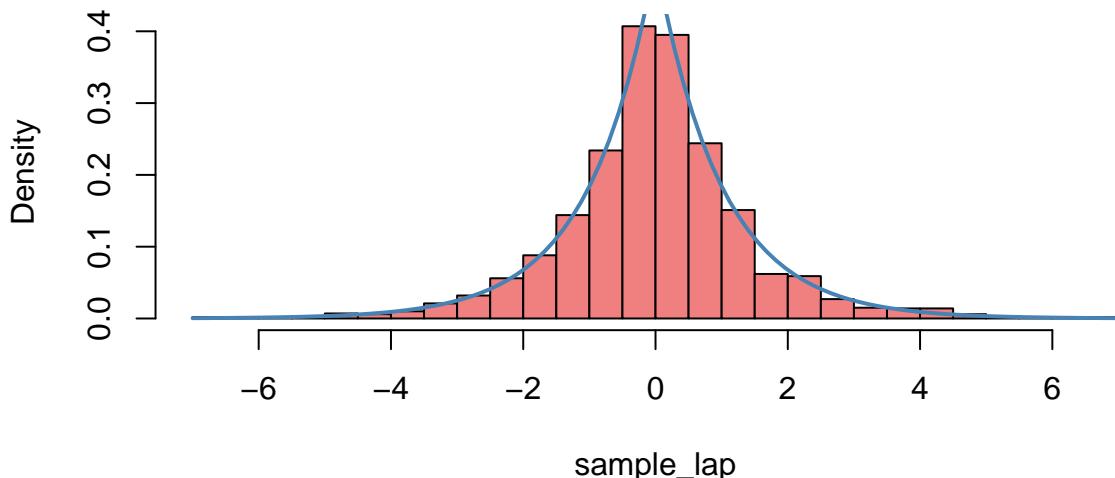


Figure 5: Rozkład Laplace'a - symulacja vs teoria.

Histogram wygenerowanej zmiennej Laplace'a pokazuje typowy szczyt w okolicy zera oraz cięjsze ogony w porównaniu z rozkładem normalnym. Parametry dobrze odwzorowują wartości wynikające z zastosowanego parametru λ , co świadczy o poprawnym działaniu generatora.

3 Generowanie rozkładu normalnego metodą Boxa-Müllera

Jeżeli U_1, U_2 są dwiema niezależnymi zmiennymi z rozkładu jednostajnego na $[0, 1]$, to zmienne

$$X_1 := \sqrt{-2 \ln(U_2)} \cos(2\pi U_1), \quad X_2 := \sqrt{-2 \ln(U_2)} \sin(2\pi U_1),$$

są dwiema niezależnymi zmiennymi o standardowym rozkładzie normalnym.

3.1 Funkcja realizująca generator par liczb losowych z rozkładu normalnego (dla zadanych parametrów μ i σ) wykorzystujący powyższą transformację Box-M'ullera.

```
rnorm_boxmuller <- function(n, mu = 0, sigma = 1){
  if(n <= 0) stop("n musi być dodatnie")
  n_pairs <- ceiling(n / 2)
  u1 <- runif(n_pairs)
  u2 <- runif(n_pairs)

  r <- sqrt(-2 * log(u2))
  theta <- 2 * pi * u1

  x1 <- r * cos(theta)
  x2 <- r * sin(theta)

  y1 <- mu + sigma * x1
  y2 <- mu + sigma * x2

  out <- c(y1, y2)
  return(out[1:n])
}
```

3.2 Histogram z nałożoną funkcją gęstości dla próby rozmiaru $n = 5000$ wygenerowanej przy pomocy generatora.

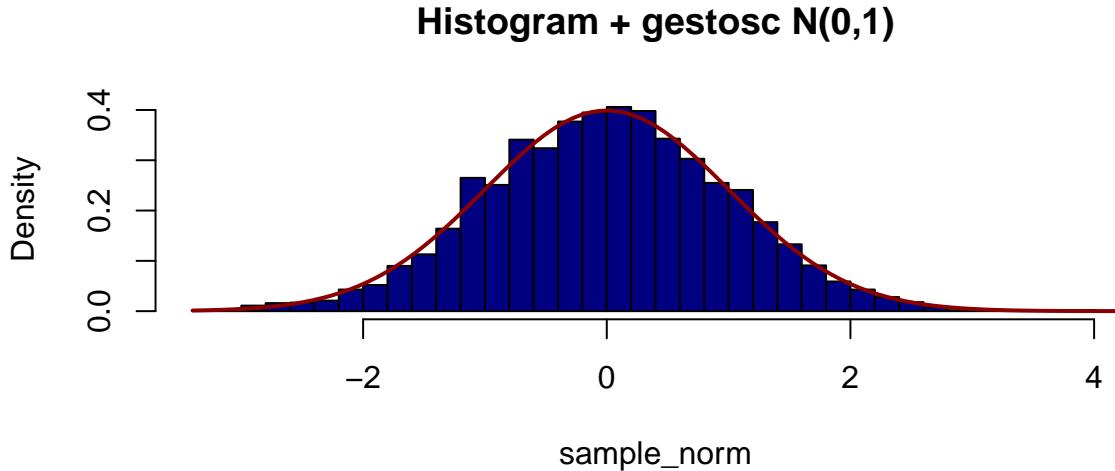


Figure 6: Histogram rozkładu normalnego z metody Boxa-Müllera.

Histogram wygenerowanych wartości jest bardzo zbliżony do krzywej gęstości rozkładu normalnego, co świadczy o poprawnym działaniu transformacji Boxa–Müllera.

3.3 Generowanie 2000 realizacji par (X_1, X_2) funkcją:

```
jointplot <- function(x,y){
  df <- data.frame(x,y)
  scatter <- ggplot(df, aes(x=x,y=y))+geom_bin2d() +
    scale_fill_continuous(type = "viridis") +
    theme_bw()
  hist_right <- ggplot(df)+geom_histogram(aes(y), fill="#69b3a2",
                                             color='darkblue')+
    coord_flip()
  hist_top <- ggplot(df)+geom_histogram(aes(x), fill="#69b3a2",
                                         color='darkblue')

  arrangeGrob(hist_top, scatter, hist_right, ncol=2, nrow=2,
              widths=c(4, 1), heights=c(1, 4),
              layout_matrix = rbind(c(1, NA), c(2,3)))}
```

- X_1 i X_2 ,
- $X_1, X_1 + X_2$,
- $X_1 + X_2, X_1 - X_2$.

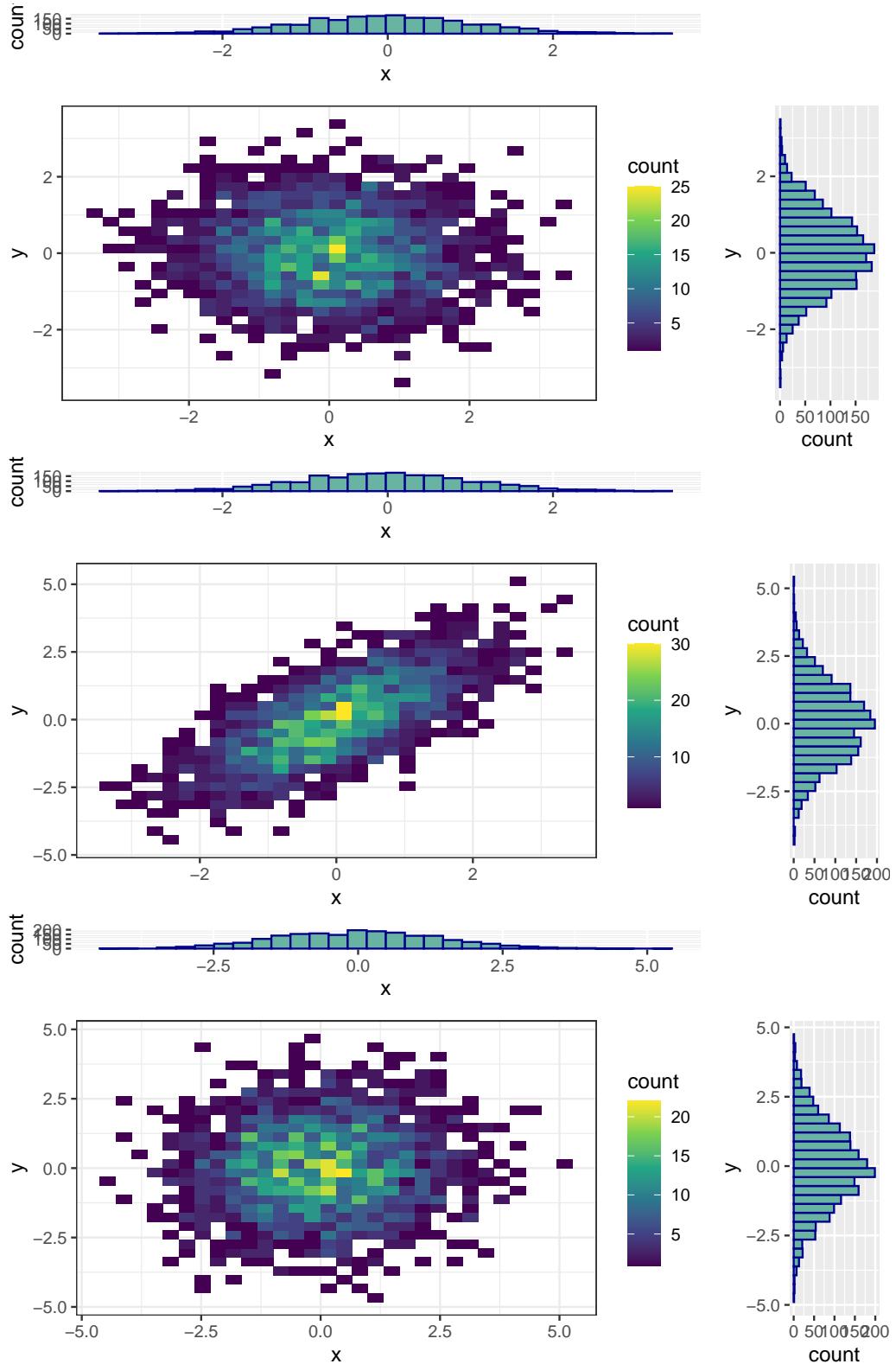


Figure 7: Wykresy łączne rozkładów: (a) X_1 i X_2 , (b) X_1 i $X_1 + X_2$, (c) $X_1 + X_2$ i $X_1 - X_2$.

3.4 Obliczenie korelacji realizacji obu zmiennych + sprawdzenie niezależności.

```
## Korelacja(X1, X2): -0.04121996
## Korelacja(X1, X1+X2): 0.7114092
## Korelacja(X1+X2, X1-X2): 0.05130102
```

Korelacje:

- $\text{corr}(X_1, X_2) \approx 0$
- $\text{corr}(X_1, X_1 + X_2) \approx \frac{1}{\sqrt{2}}$
- $\text{corr}(X_1 + X_2, X_1 - X_2) \approx 0$

Czy współrzędne są niezależne?

1. Para (X_1, X_2)

Z konstrukcji Boxa–Müllera zmienne te są *ściśle niezależne*. Korelacja bliska 0 potwierdza brak zależności liniowej.

2. Para $(X_1, X_1 + X_2)$

Zmienna $X_1 + X_2$ zawiera w sobie X_1 , więc te dwie zmienne nie mogą być niezależne.

Korelacja wychodzi bliska:

$$\text{corr}(X_1, X_1 + X_2) = \frac{\text{Var}(X_1)}{\sqrt{\text{Var}(X_1)\text{Var}(X_1 + X_2)}} \approx 0.707.$$

3. Para $(X_1 + X_2, X_1 - X_2)$

Sprawdźmy kowariancję tych zmiennych, korzystając z liniowości wartości oczekiwanej i niezależności X_1, X_2 :

$$\text{Cov}(X_1 + X_2, X_1 - X_2) = \text{Var}(X_1) - \text{Cov}(X_1, X_2) + \text{Cov}(X_2, X_1) - \text{Var}(X_2).$$

Ponieważ $X_1, X_2 \sim N(0, 1)$ są niezależne, ich kowariancje to 0, a wariancje to 1. Zatem:

$$\text{Cov}(X_1 + X_2, X_1 - X_2) = 1 - 0 + 0 - 1 = 0.$$

Oznacza to niezależność. Wektor (X_1, X_2) ma łączny rozkład normalny. Transformacja jest liniowa, więc wektor wynikowy również ma łączny rozkład normalny. Kluczową własnością wielowymiarowego rozkładu normalnego jest to, że **brak korelacji (zerowa kowariancja) jest równoważny niezależności zmiennych**.

4 Metoda akceptacji

4.1 Alternatywny sposób generowania prób z rozkładu normalnego oparty na metodzie akceptacji. Jako rozkład majoryzujący rozkład normalny wybierany rozkład Laplace'a. Wyznaczanie wartości stałej normującej M oraz parametru λ rozkładu Laplace'a, tak by możliwie zminimalizować liczbę odrzucanych próbek.

Obliczenia parametrów $i M$ dla rozkładu Laplace'a w metodzie akceptacji:

Mamy gęstości:

Normalny:

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

Laplace'a:

$$g(x) = \frac{\lambda}{2} e^{-\lambda|x|}.$$

Szukamy stałej M takiej, że:

$$f(x) \leq Mg(x) \quad \text{dla wszystkich } x.$$

Badamy iloraz:

$$\frac{f(x)}{g(x)} = \frac{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}}{\frac{\lambda}{2} e^{-\lambda|x|}} = \frac{2}{\lambda\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} + \lambda|x|\right).$$

Maksymalizujemy wykładnik:

$$h(x) = -\frac{x^2}{2} + \lambda|x|.$$

Dla $x \geq 0$:

$$h'(x) = -x + \lambda.$$

Warunek maksimum:

$$-x + \lambda = 0 \quad \Rightarrow \quad x = \lambda.$$

Wartość w maksimum:

$$h(\lambda) = -\frac{\lambda^2}{2} + \lambda^2 = \frac{\lambda^2}{2}.$$

Zatem:

$$M(\lambda) = \frac{2}{\lambda\sqrt{2\pi}} \exp\left(\frac{\lambda^2}{2}\right).$$

Minimalizujemy:

$$\ln M(\lambda) = \ln \left(\frac{2}{\sqrt{2\pi}} \right) - \ln \lambda + \frac{\lambda^2}{2}.$$

Pochodna:

$$\frac{d}{d\lambda} \ln M(\lambda) = -\frac{1}{\lambda} + \lambda.$$

Warunek:

$$-\frac{1}{\lambda} + \lambda = 0 \quad \Rightarrow \quad \lambda^2 = 1.$$

Stąd optimum:

$$\boxed{\lambda = 1}.$$

Dla $\lambda = 1$:

$$M = \frac{2}{\sqrt{2\pi}} e^{1/2}.$$

Czyli:

$$\boxed{M = \frac{2}{\sqrt{2\pi}} e^{1/2}}.$$

```
rLaplace <- function(n, lambda){
  u <- runif(n) - 0.5
  x <- -(1/lambda) * sign(u) * log(1 - 2*abs(u))
  return(x)
}

lambda <- 1
M <- (2 / sqrt(2*pi)) * exp(0.5)
```

4.2 Porównanie czasu generowania próby o rozmiarze $n = 10000$ za pomocą tego generatora oraz generatora wykorzystującego transformację Boxa–Müllera.

```
## Czas metody akceptacji: 0.06332994 s
## Czas metody Boxa–Müllera: 0.002460003 s
```

Porównanie czasu działania pokazuje, że metoda Boxa–Müllera jest o wiele szybsza, co jest typowe, bo metoda akceptacji–odrzucenia wymaga odrzucania części kandydatów. Mimo to uzyskiwane wartości są również poprawne i zgodne z rozkładem normalnym.