Kaczmarek Wiktoria, Matematyka w technice, Aktywność 1.

Rozkład dyskretny

Rozkład dwumianowym z parametrami $n\epsilon\mathbb{N}$ i $p\epsilon(0,1)$

- Definicja:

Dyskretny rozkład prawdopodobieństwa opisujący liczbę sukcesów w ciągu n niezależnych prób, każda próba ma stałe prawdopodobieństwo równe p. Opisujemy to jako $X\sim bin(n,p)$.

Funkcja prawdopodobieństwa: $PMF: f_{bin}(x|n,p) = \binom{n}{x} p^x (1-p)^{n-x} 1_{\{0,1,..,n\}}(x)$, x=0,1,..,n

- Własności:

-Wartość oczekiwana EX=np.

-Wariancja D^2X=np(1-p)

-Moda $mo(X) = \begin{cases} \lfloor (n+1)p \rfloor & , \ dla\ (n+1)p \notin \mathbb{N} \\ (n+1)p,\ (n+1)p-1 & , \ dla\ (n+1)p \in \mathbb{N} \end{cases}$

-Jeżeli ciąg zmiennych losowych X_1, X_2, … , X_n o rozkładzie SRS o rozkładzie Bernoulliego, to ich suma T_n= X_1+ X_2 + … + X_n ma rozkład dwumianowy z parametrami n i p.

$$(\forall_i\ X_i \sim B(p)) \Rightarrow T_n \sim \mathrm{bin}(n,p)$$

-Jeżeli $X\sim bin(n,p)\ i\ Y\sim bin(m,p)$ są dwiema niezależnymi zmiennymi losowymi o rozkładzie dwumianowym, wtedy ich suma X+Y jest zmienną losową o rozkładzie dwumianowym

$$X+Y\sim bin(n+m,p).$$

-Jeżeli n jest bardzo duże, a p jest bardzo małe, dobrym przybliżaniem rozkładu dwumianowego jest rozkład Poissona z parametrem λ=np.

- Przykład zastosowań:

W pewnej fabryce produkuje się 2 gatunki danego produktu: 40% produkcji to wyrób I gatunku, natomiast pozostała część to wyrób II gatunku. Odbiorca planuje zamówić 5 losowo wybranych produktów, wybór jest próbą niezależną. Oblicz prawdopodobieństwo, że dokładnie 2 produkty będą z I gatunku.

X_i - zmienna losowa opisująca wylosowanie wyrobu I, bądź II gatunku, dla i=1,2,3,4,5

Dane: n=5 próby niezależne, sukces-wylosowanie wyrobu I gatunku o prawdpodobieństwie p=0,4.

Korzystając z wzoru na $PMF: f_{bin}(x|n,p) = \binom{n}{x} p^x (1-p)^{n-x} 1_{\{0,1,..,n\}}(x)$

$$f_{bin}(2|5{,}0.4) = \binom{5}{2} 0.4^2 (1-0.4)^{5-2}$$

$$f_{bin}(2|5{,}0.3) = 10 \cdot 0.16 \cdot 0{,}216 = 0{,}3456$$

Prawdopodobieństwo, że odbiorca zamawiając 5 produktów dostanie dokładnie 2 produkty I gatunku wynosi 0.1944, czyli 19,44%.

- Możliwości wspomagania komputerowego:

Mamy duże możliwości Octave, R, Matlab.

W języku R:

```
dbinom(x, size, prob, log = FALSE)
pbinom(q, size, prob, lower.tail = TRUE, log.p = FALSE)
qbinom(p, size, prob, lower.tail = TRUE, log.p = FALSE)
rbinom(n, size, prob)
```

`x, q`	vector of quantiles.
`p`	vector of probabilities.
`n`	number of observations. If `length(n) > 1`, the length is taken to be the number required.
`size`	number of trials (zero or more).
`prob`	probability of success on each trial.
`log, log.p`	logical; if TRUE, probabilities p are given as log(p).
`lower.tail`	logical; if TRUE (default), probabilities are $P[X \le x]$, otherwise, $P[X > x]$.

W Matlabie mamy również funkcje wspomagające: binocdf, bunopdf, binoinv, gdzie binocdf(x,n,p);

Rozwiązując zadanie z przykładu w R:

```
> dbinom(2,5,0.4)
[1] 0.3456
```

Możemy też zobaczyć jakie będzie prawdopodobieństwo kiedy zadamy wektor

```
> x<-0:5
> dbinom(x,5,0.4)
[1] 0.07776 0.25920 0.34560 0.23040 0.07680 0.01024
```

Rozkład ciągły

Rozkład normalny z parametrami $\mu \in \mathbb{R}\ i\ \sigma > 0$.

- Definicja: Ciągły rozkład prawdopodobieństwa, jest to rozkład symetryczny.

Rozkład normalny może być jednoznacznie zdefiniowany poprzez:

1. Funkcja gęstości prawdopodobieństwa
2. Dystrybuantę
3. Funkcję charakterystyczną
4. Momenty.

Zmienna losowa X ma rozkład N(μ, σ) oznaczamy to $X \sim \mathrm{N}(\mu, \sigma)$.

Funkcja gęstości prawdopodobieństwo: PDF: $f_N(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), x \in \mathbb{R}$

- Własności:

Wartość oczekiwana EX= μ

Odchylenie standardowe DX= σ

Wariancja $D^2X = \sigma^2$

- Przykład zastosowań:

Pewien zakład produkcyjny zatrudnia 100 pracowników, który jest zgodny z rozkładem normalnym N(10 lat, 5 lat). Oblicz ilu pracowników miało staż dłuższy niż 15lat.

X_i - zmienna losowa opisująca długość stażu w firmie, dla i=1,2,..,100

$$X \sim \mathrm{N}(10, 5)$$

Dane: Każdy pracownik pracuje średnio 10 lat, z odchylenie standardowe wynosi 5 lat.

Korzystamy z Centralnego Twierdzenia Granicznego.

$P(X<15)=P(\frac{X-10}{5} > \frac{15-10}{5})=1-\phi(1) \approx 1 - 0{,}84134 = 0{,}15866$

$0{,}15866 \cdot 100 \approx 16$

W tym zakładzie jest około 16 osób, które mają staż pracy dłuższy niż 15lat.

- Możliwości komputerowego wspomagania:

W Exelu:

ROZKŁAD.NORMALNY(x;średnia;odchylenie_std;skumulowany)

W składni funkcji ROZKŁAD.NORMALNY występują następujące argumenty:

- **X** Argument wymagany. Wartość, dla której należy obliczyć rozkład.
- **Średnia** Argument wymagany. Średnia arytmetyczna rozkładu.
- **Odchylenie_std** Argument wymagany. Odchylenie standardowe rozkładu.
- **Skumulowany** Argument wymagany. Wartość logiczna, która określa postać funkcji. Jeśli wartością argumentu „skumulowany" jest PRAWDA, funkcja ROZKŁAD.NORMALNY zwraca funkcję rozkładu skumulowanego, a jeśli FAŁSZ, funkcja zwraca funkcję masy prawdopodobieństwa.

Obliczenie do naszego przykładu:

=1-ROZKŁ.NORMALNY(15;10;5;1)

H
0,158655

W języku R:

```
dnorm(x, mean = 0, sd = 1, log = FALSE)
pnorm(q, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)
qnorm(p, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)
rnorm(n, mean = 0, sd = 1)
```

Arguments

`x, q`
vector of quantiles.

`p`
vector of probabilities.

`n`
number of observations. If `length(n) > 1`, the length is taken to be the number required.

`mean`
vector of means.

`sd`
vector of standard deviations.

`log, log.p`
logical; if TRUE, probabilities p are given as log(p).

`lower.tail`
logical; if TRUE (default), probabilities are $P[X \le x]$ otherwise, $P[X > x]$.

Obliczenie do naszego przykładu:

```
> pnorm(15,10,5,FALSE,FALSE)
[1] 0.1586553
```