Wyznaczyć E(Y) oraz D^2 (Y) dla zmiennej losowej $Y = e^{aX}$.

Rozwiązanie:

$$E(Y) = \sum_{i=0}^{n} y_i P_{n,i} = \sum_{i=0}^{n} e^{ai} \binom{n}{i} p^i q^{n-i} = \left(q + pe^{a}\right)^n$$

$$D^2Y = \sum_{i=0}^{n} y_i^2 P_{n,i} - (EY)^2 = \sum_{i=0}^{n} \binom{n}{i} e^{2ia} p^i q^{n-i} - \left(q + pe^{a}\right)^{2n} = \sum_{i=0}^{n} \binom{n}{i} \left(e^{2a} p\right)^i q^{n-i} - \left(q + pe^{a}\right)^{2n}$$

$$= \left(q + pe^{2a}\right)^n - \left(q + pe^{a}\right)^{2n}$$

4. Opracowanie statystyczne dla danych empirycznych

Najważniejsze statystyki to średnia i wariancja empiryczna.

Średnią empiryczną dla danych z próby, zwaną potocznie statystyką z kreską, która jest odpowiednikiem średniej arytmetycznej wyraża się wzorem:

(1.18) $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ dla danych niepogrupowanych;

oraz

(1.19) $\overline{X} = \frac{1}{n}\sum_{i=1}^{k} x_{is} n_i$ dla danych pogrupowanych, gdzie k jest liczbą klas; x_{is} jest środkiem i-tego przedziału klasowego.

A zatem $x_{is} = \frac{a_i + a_{i+1}}{2}$

Wariancję empiryczną zwaną potocznie statystyką „S kwadrat" obliczamy ze wzoru:

(1.20) $S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ dla danych niepogrupowanych;

lub

(1.21) $S^2 = \frac{1}{n}\sum_{i=1}^{k}\left(x_{is} - \overline{X}\right)^2 n_i$ dla danych pogrupowanych, gdzie k oraz x_{is} oznaczają to samo co we wzorze **(1.19).**

Liczbę klas k ustala się w zależności od liczby obserwacji n. Na przykład można ustalić ***k*** wg wzoru: $k \approx \sqrt{n}$.

Pozostałe parametry obliczane dla danych empirycznych będą definiowane w przykładach.

60/ *Dokonano 25 niezależnych prób otrzymując wyniki:*
0,72; 0,65; 0,91; 0,70; 0,02; 0,65; 0,76; 0,88; 0,25; 0,61; 0,12; 0,17; 0,31; 0,37; 0,41; 0,17; 0,28; 0,66; 0,11; 0,10; 0,22; 0,23; 0,17; 0,19; 0,26. *Obliczyć średnią i wariancję empiryczną.*

Rozwiązanie:

Średnią i wariancję empiryczną obliczamy ze wzorów: (1.18) oraz (1.20).

$$\overline{X} = \frac{1}{25}\sum_{i=1}^{25} x_i = \frac{1}{25}\cdot(9{,}92) = 0{,}3968 \approx 0{,}4$$

$$s^2 = \frac{1}{25}\sum_{i=1}^{25}\left(x_i - \overline{X}\right)^2 = \frac{1}{25}\begin{bmatrix}(0{,}72-0{,}4)^2 + (0{,}65-0{,}4)^2 + (0{,}91-0{,}4)^2 + (0{,}7-0{,}4)^2 + \ldots \\ \ldots + (0{,}19-0{,}4)^2 + (0{,}26-0{,}4)^2\end{bmatrix} = 0{,}071$$

(Przy obliczaniu wariancji pokazano tylko fragment obliczeń aby przedstawić sposób rozpisywania wzoru)

61/ *Zmienna losowa X ma rozkład jednostajny na odcinku $<a,a+1>$. Wykonano n niezależnych obserwacji tej zmiennej losowej. Sprawdzić, że T_n jest estymatorem zgodnym i nieobciążonym dla parametru a, jeżeli:*

$$T_n = \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{2}$$

Rozwiązanie:

Estymator parametru a - oszacowanie statystyczne parametru a;

Mówimy, że **estymator jest zgodny** gdy: $\underset{\varepsilon>0}{\forall} \lim_{n\to\infty} P(|T_n - a| < \varepsilon) = 1$ **(1.22)**

Mówimy, że **estymator jest nieobciążony**, gdy $ET_n = a$. **(1.23)**

gdzie ***a*** jest szacowanym parametrem.

Dla zmiennej losowej X gęstość prawd. można zapisać:

$$f(x) = \begin{cases} 1 & dla \quad x \in <a,a+1> \\ 0 & dla \quad x \notin <a,a+1> \end{cases}$$

$$EX = \int_a^{a+1} x\,dx = \left[\frac{x^2}{2}\right]_a^{a+1} = \frac{(a+1)^2}{2} - \frac{a^2}{2} = \frac{2a+1}{2} = a + \frac{1}{2}$$

$$ET_n = \frac{1}{n}\left(n\left(a + \frac{1}{2}\right)\right) - \frac{1}{2} = a$$

A zatem T_n jest estymatorem nieobciążonym dla parametru a.

Sprawdźmy jeszcze czy T_n jest estymatorem zgodnym. Aby tak było musi być spełniony warunek, dla $n \to \infty$, wariancja $D^2T_n \to 0$.

(Wynika to ze słabego prawa wielkich liczb dla ciągu niezależnych zmiennych los.)

Pomocniczo obliczamy wariancję D^2X.

$$E(X^2) = \int_a^{a+1} x^2 dx = \left[\frac{x^3}{3}\right]_a^{a+1} = \frac{(a+1)^3}{3} - \frac{a^3}{3} = a^2 + a + \frac{1}{3}$$

$$D^2X = E\left(X^2\right) - (EX)^2 = a^2 + a + \frac{1}{3} - \left(a + \frac{1}{2}\right)^2 = \frac{1}{12}$$

A zatem wariancja dla T_n wynosi: