1. Nacrtati histogram i boxplot za sledeće podatke. Prokomentarisati rezultate.

1.7, 0.8, 1.2, 0.8, 0.7, 0.2, 0.2, 0.2, 2.4, 0.4, 0.7, 2.9, 1.1, 0.5, 1.6, 0.2, 0.2, 0.2, 0.1, 1.0, 2.4, 1.2, 0.2, 0.3, 0.8

Rešenje:

Odredimo ponovo varijacioni niz:

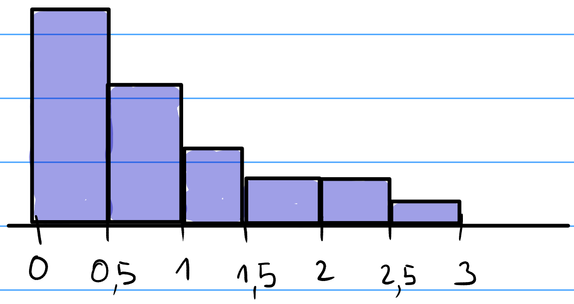
0.1, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.3, 0.4, 0.5, 0.7, 0.7, 0.8, 0.8, 0.8, 1.0, 1.1, 1.2, 1.2, 1.6, 1.7, 2.4, 2.4, 2.9

$n = 25$, $R = x_{(n)} - x_{(1)} = 2{,}9 - 0{,}1 = 2{,}8$, $k = \lceil \log_2 n \rceil + 1 = 6$

$$h = \frac{R}{k} = 0{,}467 \approx 0{,}5$$

Najmanja tačka uzorka je 0.1, pa ćemo za početnu tačku prvog intervala histograma uzeti 0. Dakle, intervali nad kojima crtamo stupce histograma su:

$$\underbrace{(0, 0{,}5]}_{\text{ovde je 11 tačaka}}, \underbrace{(0{,}5, 1]}_{\text{-||- 6 tačaka}}, \underbrace{(1, 1{,}5]}_{\text{-||- 3 tačke}}, \underbrace{(1{,}5, 2]}_{\text{-||- 2 tačke}}, \underbrace{(2, 2{,}5]}_{\text{-||- 2 tačke}}, \underbrace{(2{,}5, 3]}_{\text{1 tačka}}$$



visine ovih stubaca su redom 11,6,3,2,2,1

Ovo je kao i u prethodnom primeru bio histogram frekevncija. Iz histograma frekvencija možemo takođe pretpostaviti gustinu raspodele, koja bi u ovom slučaju mogla biti eksponencijalna. Ako baš želimo da ocenimo gustinu ovim "stepenicama", potrebno je da visinu svakog stupca sa nh. Oblik histograma se tako ne menja, ali dobija se lepo svojstvo, a to je da je ukupna površina ispod stubaca jednaka 1, a 1 je ukupna površina ispod grafika gustine.
Ovo važi jer je u histogramu gustine površina svakog stupca jednaka:

$$h \cdot \frac{\text{br. tačaka koje su se našle u intervalu}}{nh},$$

pa kad saberemo sve površine, dobijamo:

$$\frac{\text{br. tačaka u 1. intervalu}}{h} + \dots + \frac{\text{br. tačaka u k. intervalu}}{h} = \frac{\text{ukupan br. tačaka}}{h} = \frac{n}{n} = 1$$

Parametar eksponencijalne raspodele možemo oceniti iz podataka kao u nekom od naših zadataka:

$$\hat{\lambda} = \frac{1}{\bar{X}_n}$$

Za crtanje boxplota određujemo uzoračke kvartile:

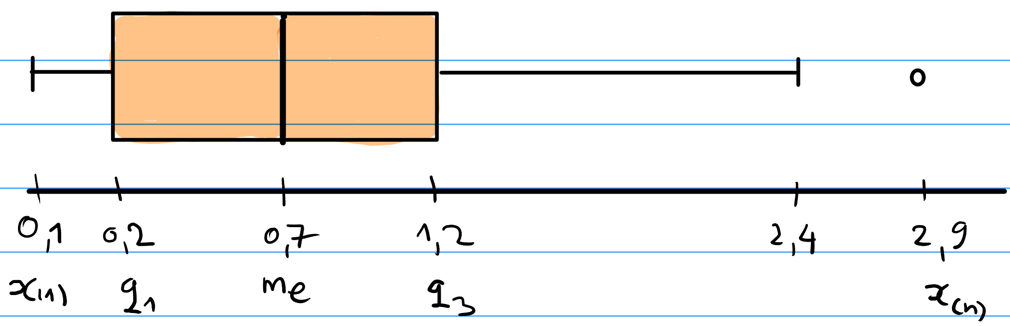
$me = 0{,}7 \qquad q_1 = 0{,}2 \qquad q_3 = 1{,}2 \qquad IQR = q_3 - q_1 = 1$

$q_1 - 1{,}5\, IQR = -1{,}3 < 0{,}1 = x_{(1)} \qquad q_3 + 1{,}5\, IQR = 2{,}7 < 2{,}9 = x_{(n)}$

nema autlajera sa ove strane

dakle, imamo 1 autlajer i to je najveća tačka uzorka

granice boxplota su: $0{,}1$ i $\underbrace{2{,}4}_{\text{najveća tačka uzorka koja nije veća od 2.7}}$



Vidimo da je ova kutija pomerena ka levom kraju boxplota, pa zaključujemo da raspodela nije simetrična. Takođe, vidimo da se tačke skoncentrišu na levom kraju, a da su na desnom prilično razuđene. (25% tačaka je i u [0.1,0.2] i u [1.2,2.9]).