Funkcja sd wylicza estymator nieobciążony (n-1).

Zad 1

Miary położenia:

* średnia
* mediana
* moda
* kwartyle q1 i q3

Miary rozproszenia

* zakres
* wariancje i odchylenie standardowe nieobciążone (s\*^2 i s\*) oraz wariancje i odchylenie standardowe obciążone (s^2 i s)
* odchylenie od wartości przeciętnej d1
* odchylenie przeciętne od mediany d2
* odchylenie ćwiartkowe Q
* współczynnik zmienności V
* pozycyjny współczynnik zmienności VQ

Miary asymetrii

* skośność g1
* kurtoza krt
* eksces g2

Jak znaleźć szereg rozdzielczy?

Do wyznaczenia szeregu można użyć funkcji hist. Spełnia ona 2 cele, może wykreślić standardowo histogram, ale dane do tego histogramu są również utworzone. Gdy nacisniemy na liste h1 to pojawi się lista elementów tej struktury. Co oznacza breaks : punkty podziału

Counts: liczba zliczonych obserwacji (dla szeregu rozdzielczego)

Mids: Środki przedziałów, można każdą statystykę na podstawie szeregu rozdzielczego znaleźć.

Gdyby ktoś miał średnią policzyć z tego zbioru na podstawie tego szeregu rozdzielczego to dla szeregu rozdzielczego nie możemy zastosować funkcji mean(), która liczy średnią bo jest dla szeregu szczegółowego, ale można wziąć sumę iloczynów podzielić przez liczbę obserwacji. Funkcja sum(iloczyn tej wielkości liczby obserwacji \* środki i to wszystko podzielić przez sumę zliczonych obserwacji)

Jeżeli chodzi o odnoszenie się do pól, to nazwa struktury h1, elementem rozdzielającym listę od pola jest $ czyli h1$counts \* h1$mids to powinno być pod znakiem sumowania przez sumę h1$counts.

W przypadku wielkości opisowych typu

* kwartyle
* mediana
* moda

proponuję by do wyliczenia tych wielkości z szeregu rozdzielczego skorzystać ze zbioru zadan. Rozdział 4, zadanie 2 pokazuje odpowiednie wzory na wyliczenie tych wielkości. Żeby te sumy zastosować, należy wiedzieć, w którym przedziale w których dana wielkość wystepuje. Do wyznaczenia przedziałów kwartyli i mediany wygodna jest funkcja findInterval(), która znajdzie przedział, gdzie odpowiedni kwartyl będzie umieszczony minus numer przydziału. W przypadku mody w większości zadań moda dla szeregu może nie istnieć, ale w rozdzielczym może występować. Funkcja which.max() wyznaczy przedział, gdzie moda się znajduje.

zad 2

Rozdział 6.2, zadanie 5

zad o numerach nieparzystych i dla zadania 5 na wykladzie były wskazane 2 sposoby realizacji testu, 1 z wykorzystaniem zbioru krytycznego (podałem jakie funkcje mozna te kwartyle mozna wykorzystywać).

2 sposób polega na wyznaczeniu wartość ..., czyli gdy prawdopobienstwo jest bardzo małe albo bardzo duze, to oznacza że (14:02).

help(runif) - wyjasnienie od 14:20

Jak

Zad 5

Celem jest porównanie dwóch wartości przeciętnych. Należy użyć test na równość wartości przeciętnych. Trzeba sprawdzić gdzieś, którą statystykę użyć, do wyboru statystyki może być pomocna informacja czy wariancje są identyczne czy różne. Test wartości przeciętnych poprzedza się testem równości wariancji (25:36). W zależności od tego wyniku tego testu wybiera się odpowiednią statystykę do porównania wartości przeciętnych.