MWS LAB5

Piotr Mikołajczyk

12.06.2021

Zad1

a) Oszacuj wartości oczekiwane u1 i u2 rozkładów, z których generowane były liczby oraz ich różnicę u1 - u2.

Wartości oczekiwane u1 = 1.017 ; u2 = 1.204

Różnica średnich wynosi : 0.187

b) Oszacuj wariancję sigma2.

$$s^2 = \frac{(n-1)s_X^2 + (m-1)s_Y^2}{m+n-2}, \quad s_X^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}, s_Y^2 = \frac{\sum_{i=1}^{m}(X_i - \overline{Y})^2}{m-1}.$$

Wzór na wariancje sigma2

Gdzie: n – ilość liczb wylosowanych przez generator 1, m – ilość liczb wylosowanych przez generator 2, sx^2 – wariancja liczb wylosowanych przez generator 1, sy^2 -wariancja liczb wylosowanych przez generator 2

Wariancja sigma2 wynosi: 0.960092

c) Oszacuj odchylenie standardowe błędu wy estymowanej w punkcie a. różnicy u1 - u2.

$$s = \sqrt{s^2\left(\frac{1}{n} + \frac{1}{m}\right)}$$

Wzór na odchylenie standardowe błędu różnicy u1 i u2

Odchylenie standardowe jest równe: 0.6572986

d) Czy w opisanej sytuacji właściwym testem do sprawdzenia równości między średnimi będzie test jednostronny, czy dwustronny?

Dla testu jednostronnego u1 =/= u2 -; Dla testu dwustronnego u1 = u2 W tym wypadku w tej sytuacji lepszy będzie test dwustronny

e) Czy hipoteza o równości średnich zostałaby odrzucona przez dwustronny test na poziomie istotności alfa = 0:1? Nie, ponieważ wartość hipotezy 0 jest mniejsza od p-wartości. Poniżej wzory na wyznaczenie granicy decyzyjnej oraz wartości progowej.

$$T = \frac{\overline{X} - \overline{Y}}{s\sqrt{n^{-1} + m^{-1}}}.$$

$$|T| > c, \qquad c = F^{-1}_{t_{m+n-2}}\left(1 - \frac{\alpha}{2}\right)$$

dev : 0.6572986 ; p-value: 0.7843

Zad2

a.) Przeprowadź test tego, że nie ma różnicy między łożyskami wykonanymi z tych materiałów, zakładając, że czas pracy do momentu uszkodzenia opisuje się rozkładem normalnym.

Ponieważ zakładamy rozkład normalny przeprowadzić można test studenta który potwierdził że nie ma różnicy między łożyskami wykonanymi z tych materiałów.

b) Przeprowadź analogiczny test, bez zakładania normalności rozkładów.

Test Wilcoxona dał przybliżony wynik p-wartości, wobec czego należy przyjąć że między łożyskami nie ma różnicy.

c) Który z powyższych testów jest w rozważanym przypadku odpowiedniejszy?

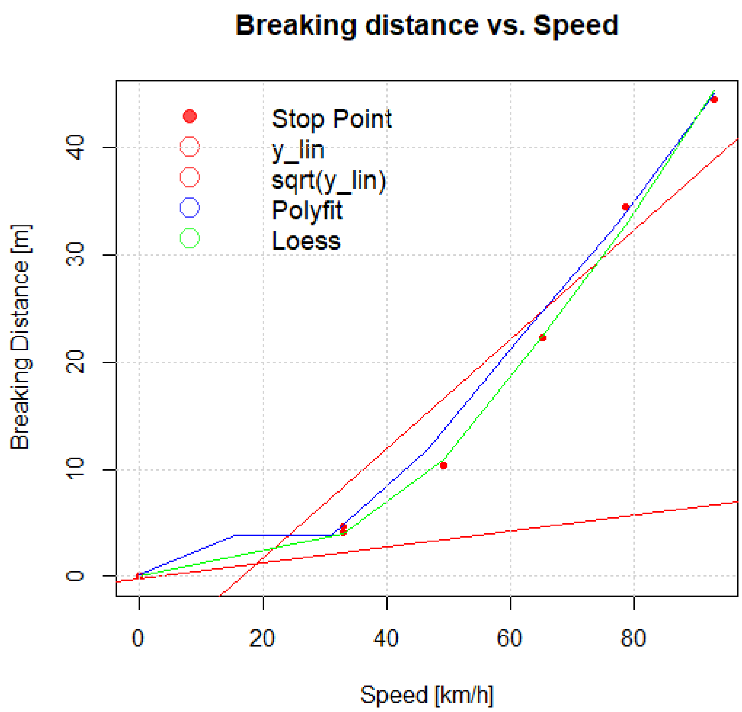
Odpowiedniejszym testem jest test Wilcoxona ponieważ w podpunkcie 2a jedynie zakładamy że czas pracy do momentu uszkodzenia opisuje się rozkładem normalnym co w istocie może nie być prawdą.

Zad3

W poniższej tabeli przedstawiona jest długość drogi hamowania y (pewnego pojazdu, na pewnego rodzaju drodze) w funkcji prędkości v. Dopasuj liniowe zależności opisujące długość drogi y oraz pierwiastek tej długości sqrt(y) w funkcji prędkości v. Która z liniowych zależności jest bardziej zgodna z danymi? Dlaczego?

Tabela 1 – Dane długości drogi hamowania w funkcji prędkości pewnego pojazdu.

y [km/h]	0	33	33	49.1	65.2	78.5	93
y [m]	0	4.7	4.1	10.3	22.3	34.4	44.4



Rys. 1 – Dane z Tabeli 1 oraz regresje liniowa, pierwiastkowa, funkcja Polyfit oraz Loess dla porównania dopasowania.

Wnioski:
Bardziej zgodna z danymi jest linia pierwsza (y nie pod pierwiastkiem) ponieważ jest dobrą reprezentacją trendu funkcji drogi hamowania od prędkości. Trend ten jest kwadratowy ponieważ energia kinetyczna rośnie z kwadratem prędkości, natomiast siła tarcia zatrzymuje rozpędzoną masę liniowo siłą tarcia, co oznacza wprost że trend długości hamowania w funkcji prędkości powinien być kwadratowy. Przyjęcie pierwiastka dla y oznacza nie spełnienie fizycznych ograniczeń dotyczących kreacji tego typu danych. Z tego również powodu funkcja lokalnej regresji wielomianowej idealnie nadaje się do aproksymacji trendu ponieważ potrafi wpasować się w kwadratową naturę wykresu.
Bibliografia:

[1] Slajdy wykładowe do przedmiotu MWS, Rafał Rytel-Andrianik