

2 Model statystyczny

Jeżeli próbka $\mathbf{x} = (x_1, \dots, x_n)'$ jest **reprezentatywna**, to stanowi ona podstawę do wnioskowania o **populacji** z której pochodzi. Wnioskowanie takie wymaga zbudowania modelu “zachowania się” zmiennej (cechy) X w populacji. Budowa modelu polega na przyjęciu założenia o rozkładzie (teoretycznym) zmiennej X w populacji oraz traktowaniu obserwacji jako wartości tej zmiennej.

Dokładniej: budując model statystyczny traktujemy wektor obserwacji (próbkę) $\mathbf{x} = (x_1, \dots, x_n)'$ jak realizację wektora losowego (próby) $\mathbf{X} = (X_1, \dots, X_n)'$ z nieznanego (lub jedynie częściowo nieznanego) rozkładu.

Statystyka, to każda (mierzalna) funkcja próby. Zatem w modelu, statystyka jest wielkością losową.

Model: jedna próba prosta

Modelujemy wyniki doświadczenia w którym dokonujemy n -**niezależnych** obserwacji badanej cechy X na losowo wybranych z populacji (jednorodnej ze względu na badaną cechę) jednostkach eksperymentalnych.

Przykład 1. W celu określenia czasu bezawaryjnej pracy urządzeń po wykonaniu kapitalnego remontu, wybrano 50 urządzeń i obserwowano czas ich bezawaryjnej pracy. Wyniki (w h.) są następujące: 629, 325, 215, ... , 612, 841, 492.

Budując model statystyczny tego eksperymentu zakładamy, że (w populacji) czas bezawaryjnej pracy urządzenia (cecha X) ma rozkład wykładniczy z nieznanym parametrem λ .

Model ten ma jeden parametr: λ .

Przykład 2. Przeprowadzono 50 niezależnych eksperymentów polegających na hamowaniu badanego typu samochodu wyposażonego w nowy typ układu hamulcowego (na suchym asfalcie, przy prędkości 40 km/h, itd.). Notowano długość drogi hamowania z dokładnością do jednego centymetra. Otrzymane wyniki to: 18.66, 17.81, 18.96, ... , 17.62, 18.61, 17.99.

Budując model statystyczny tego eksperymentu zakładamy, że (w populacji) długość drogi hamowania (cecha X) ma rozkład normalny z nieznanymi parametrami μ i σ^2

Uwaga! Model ten często zapisujemy w następującej postaci:

$$X_i = \mu + \varepsilon_i, \quad i = 1, \dots, n,$$

gdzie

X_i - i -ta obserwacja badanej cechy X ,

μ - wartość oczekiwana (średnia, “prawdziwa” wartość) badanej cechy X ,

ε_i - błędy (reszty) - niezależne zmienne losowe o jednakowym rozkładzie $N(0, \sigma^2)$.

Model ten ma dwa parametry: μ i σ^2 .