

## Kísérlettervezés (kiegészítő anyag)

### Adatelemzés – elméleti ismeretek

#### Valószínűségszámítási alapfogalmak

- Valószínűségi változó (*random variable*):  $X$
- Várható érték, átlag (*expected value, average, mean*):

$$\mu = \mathbb{E}X = \sum_{i=1}^n p_i x_i$$

- Szórásnégyzet (*variance*):

$$\sigma^2 = \mathbb{E}(X - \mu)^2 = \sum_{i=1}^n p_i (x_i - \mu)^2$$

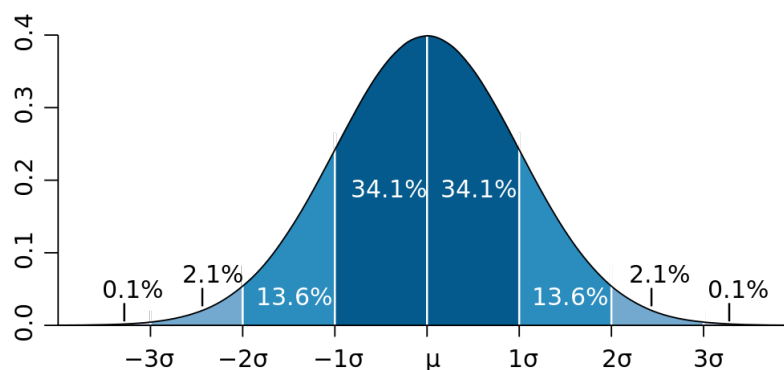
- Szórás (*standard deviation*):

$$\sigma = \sqrt{\mathbb{E}(X - \mu)^2} = \sqrt{\sum_{i=1}^n p_i (x_i - \mu)^2}$$

#### Konfidencia – normális eloszlás

A normális eloszlású változó

- az esetek 68%-ában legfeljebb  $1\sigma$  messze kerül  $\mu$ -től,
- az esetek 95%-ában legfeljebb  $2\sigma$  messze kerül  $\mu$ -től,
- az esetek 99,7%-ában legfeljebb  $3\sigma$  messze kerül  $\mu$ -től.



1. ábra. Konfidenciaintervallumok

#### Statisztikai alapfogalmak

- Megfigyelések:  $t$  darab,  $x_1, \dots, x_t$
- Tapasztalati átlag (*sample mean*):

$$m = \bar{x} = \frac{x_1 + \dots + x_t}{t}$$

- Korrigált tapasztalati szórás (*unbiased sample standard deviation*):

$$s = \sqrt{\frac{(x_1 - m)^2 + \dots + (x_t - m)^2}{t - 1}} = \sqrt{\frac{\sum_{i=1}^t (x_i - m)^2}{t - 1}}$$

Figyeljük meg, hogy a korrigált tapasztalati értékeknél  $t$  helyett  $(t - 1)$ -gyel osztunk. Ennek oka, hogy  $t$ -vel osztva a kapott érték általában alábecsli a teljes populáció szórását. Belátható, hogy  $(t - 1)$ -gyel osztva a valódi szórást jobban közelítő értéket kapunk. Ezt nevezzük Bessel-féle korrekciónak ([https://en.wikipedia.org/wiki/Bessel's\\_correction](https://en.wikipedia.org/wiki/Bessel's_correction)).

## Kísérlettervezés

A centrális határeloszlás-tételből (*central limit theorem*) következik, hogy tetszőleges eloszlású jellemző (véges  $\mu$  várható értékkel és  $\sigma$  szórással) tapasztalati átlaga ( $m$  valószínűségi változó)  $t \rightarrow \infty$  esetén normális eloszlású,  $\mu$  várható értékkel és  $\sigma_m = \frac{\sigma}{\sqrt{t}}$  szórással (ahol  $\sigma$  a megfigyelés alatt álló valószínűségi változó szórása).

Ökölszabály: ismert szórásnál  $t > 30$  után kezd elfogadható lenni ez a közelítés.

A fentiek miatt a megfigyeléseink kiértékeléséhez használhatjuk a normális eloszláshoz tartozó konfidenciaintervallumokat. Eszerint  $m$  68%-os valószínűséggel maximum  $\sigma_m$  távolságra esik  $\mu$ -tól. Tehát (ha  $m$ -et és  $\sigma_m$ -et ismerjük, és  $\mu$ -t szeretnénk ez alapján becsülni)  $\mu$  értékéről kijelenthetjük, hogy

- 68% valószínűséggel  $m - \sigma_m$  és  $m + \sigma_m$  közé esik, valamint ehhez hasonlóan
- 95% valószínűséggel  $m - 2\sigma_m$  és  $m + 2\sigma_m$  közé esik,
- és 99,7% valószínűséggel  $m - 3\sigma_m$  és  $m + 3\sigma_m$  közé esik.

Sokszor azonban  $\sigma$  nem ismert (a mérésekre azért van szükség, hogy becsléseket kaphassunk  $\mu$  és  $\sigma$  értékére), ekkor elfogadható a  $\sigma_m \approx \frac{s}{\sqrt{t}}$  (ahol  $s$  a korrigált tapasztalati szórás), amennyiben  $t \geq 100$ .

## 1. Kísérlet kiértékelése

Infrastruktúránk méretezését megnehezíti, hogy egy adott feladattípus végrehajtási ideje a körülmények függvényében ingadozik, például lapozás, memória szemétygyűjtés, memória cache találatok stb. változékonysága folytán. Ezért összeállítottunk egy valós munkaterhelést jól jellemző benchmarkot, és ennek többszöri lefuttatása során a futási időket átlagolva szeretnénk meghatározni a rendszer átlagos teljesítményét.

- a) Az első tíz futtatás eredményei: 37 s, 34 s, 35 s, 39 s, 57 s, 41 s, 36 s, 35 s, 61 s, 35 s. Mennyi ez alapján a rövid kísérlet alapján a tapasztalati átlag és tapasztalati szórás?
- b) Nagyobb léptékben futtatva a kísérletet, a benchmark 10 000 futtatása átlagban 44,3 másodpercig tartott, 11,6 másodperc tapasztalati szórással. Mennyire lehetünk biztosak a kapott eredmény pontosságában?

## 2. Kísérlettervezés

Egy modellezett folyamat átbecsátóképességére szimuláció alapján szeretnénk egy közelítő értéket és hozzá tartozó konfidencia-intervallumot meghatározni.

- a) Hány szimuláció mérési eredményeiből számoljunk átlagot?
- b) Az így elvégzett mérési eredmények tapasztalati közepe 500 kérés/s; a tapasztalati szórás 10%. Szeretnénk, hogy 95% konfidencia mellett egy legfeljebb 40 kérés/s széles intervallumba essen az átbecsátóképesség. Hány mérést végezzünk még?