Kísérlettervezés

Valószínűségszámítási alapfogalmak

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

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

- Szórásnégyzet, variancia (*variance*):

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

- Szórás (*standard deviation*):

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

Statisztikai alapfogalmak

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

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

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

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

- 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[1].

Kísérlettervezés

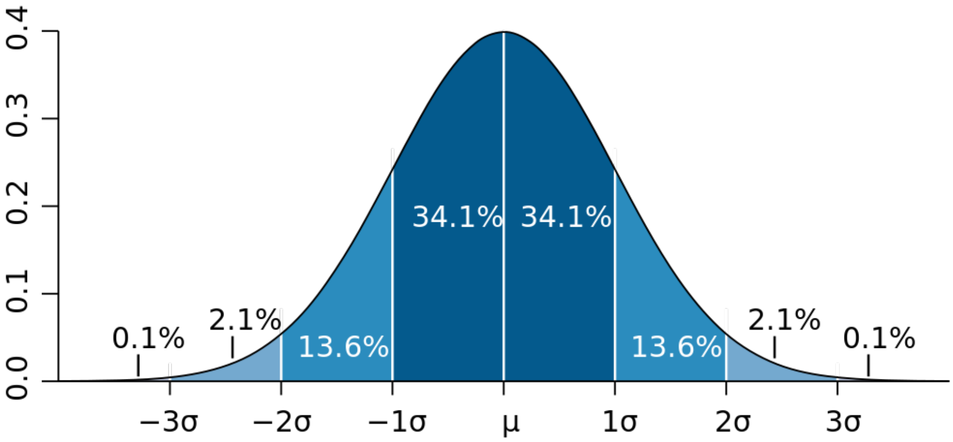
- A centrális határeloszlás tételéből (CHT) következőik, hogy tetszőleges eloszlású jellemző (véges m várható értékkel és s szórással) tapasztalati átlaga $t \to \infty$ esetén normális eloszlású, $\mu = m$ várható értékkel és $\sigma = \frac{s}{\sqrt{t}}$ szórással.

  Ökölszabály: ismert szórásnál $t > 30$, ismeretlen szórásnál $t > 100$ után kezd elfogadható lenni a közelítés.

- A normális eloszlású változó

[1] http://hu.wikipedia.org/wiki/Bessel-f%C3%A9le_korrekci%C3%B3

- az esetek 68%-ában legfeljebb 1σ messze kerül μ-től,
- az esetek 95%-ában legfeljebb 2σ messze kerül μ-től,
- az esetek 99,7%-ában legfeljebb 3σ messze kerül μ-től.



1. ábra. Konfidenciaintervallumok