FYS4411, MARCH 10, 2022

statistical error:

$$\mathbb{E}[f(x)] = \int_{x\in D} p(x) f(x)\, dx = \mu_f$$

$$\mathbb{E}[f(x)^n] = \int_{x\in D} p(x) f^n(x)\, dx$$

$$\mathrm{var}[f(x)] = \sigma_f^2 =$$

$$\mathbb{E}[f(x)^2] - \left(\mathbb{E}[f(x)]\right)^2$$

$$= \int_{x\in D} p(x) \left(f(x) - \mu_f\right)^2 dx$$

$$\simeq \frac{1}{M} \sum_{i=1}^{M} \left(f(x_i) - \bar{\mu}_f\right)^2$$

sample variance $= \bar{\sigma}_f^2$

$$\bar{\mu}_f = \frac{1}{M} \sum_{i=1}^{M} f(x_i)$$

sample mean

$$\boxed{\bar{\mu}_f \neq \mu_f} \quad \left[\bar{\sigma}_f^2 \neq \sigma_f^2\right]$$

Central limit theorem

iid (independent and identically distributed stochastic variables)

if $x \in$ iid then in the limit $n \to \infty$, the final PDF is going to be a gaussian

x given by $p(x)$

$$\bar{x} = \mu_x = \mu = \int p(x)\, x\, dx$$

$$\left(\sum_{i=1}^{n} p(x_i)\, x_i \right)$$

$$\approx \frac{1}{n} \sum_{i=1}^{n} x_i$$

Experiment x_α

$$\bar{x}_\alpha = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad \left(\sum_{i=1}^{n} p(x_i)\, x_i \right)$$

$$= \mu \qquad \int x\, p(x)\, dx$$

new variable

$$\bar{z} = \frac{x_1 + x_2 + \bar{x}_3 + \dots x_m}{m}$$

$\bar{x}_i$ are given by $p(x)$

what is the distribution which describes $\bar{z}$?

$\bar{z} \to z \qquad \bar{x}_\alpha \to x_\alpha$

$$\bar{z} = \frac{1}{m} \sum_{\alpha=1}^{m} \bar{x}_\alpha = \frac{1}{m} \frac{1}{n} \sum_{i=1}^{n} \sum_{\alpha=1}^{m} x_i$$

$$p(z) = \int dx_1 \, p(x_1) \int dx_2 \, p(x_2) \dots \int dx_m \, p(x_m) \times \delta\left(z - \frac{x_1 + x_2 + \dots + x_m}{m}\right)$$

$$\Big(p(z) = p(x_1, x_2, \dots x_m) = p(x_1)\, dx_1 \dots p(x_m)\, dx_m$$

$$\delta\left(z - \frac{x_1 + x_2 + \dots + x_m}{m}\right)$$

$$\frac{1}{} \int^{\infty} dq \, e^{iq\left(z - \frac{x_1 + x_2 + \dots + x_m}{m}\right)}$$

$$= \frac{\;}{2\pi} \int_{-\infty} \quad + \;-$$

$$p(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(iq(z-\mu))\,dq$$

$$\times \left[\int_{-\infty}^{\infty} dx\, p(x) \exp\left(iq\left(\frac{\mu - x}{m}\right)\right) \right]^m$$

$$\int_{-\infty}^{\infty} dx\, p(x) \exp\left(iq\left(\frac{\mu - x}{m}\right)\right)$$

$$= \int_{-\infty}^{\infty} dx\, p(x) \left[1 + iq\frac{(\mu - x)}{m} - q^2 \frac{(\mu - x)^2}{2m^2} + \dots \right]$$

$$\int_{-\infty}^{\infty} dx\, p(x) = 1 \quad \wedge \quad \mu = \int_{-\infty} dx\, p(x)\, x$$

$$= 1 + 0 - \frac{q^2 \sigma^2}{2m^2} + \dots$$

2

$$\sigma^2 = \int p(x)\, dx\, (x-\mu)^{-}$$

$$\left[\int_{-\infty}^{\infty} dx\, p(x)\, \exp\left(iq \frac{(\mu - x)}{m} \right) \right]^2$$

$$\simeq \left[1 - \frac{q^2 \sigma^2}{2m^2} \right]^m$$

in the limit $m \to \infty$

$$p(z) = \frac{1}{\sqrt{2\pi\sigma^2/m}} \exp\left[\frac{-(z-\mu)^2}{2(\sigma/\sqrt{m})^2} \right]$$

$$z = \frac{\bar{x}_1 + \vec{x}_2 + \cdots \bar{x}_m}{m} = \frac{m\mu}{m}$$

$$= \mu$$

new distribution (Gaussian) with mean value $= \mu$ and variance σ^2/m

The variance of x is σ^2

$$\sigma^2 = \int p(x)\, (x-\mu)^2\, dx$$

Bootstrap (resampling method)

$$X = \{x_1, x_2, x_3, \ldots x_n\}$$

algorithm

- Draw a Bootstrap sample

  $X_1 = \{x_1^*, x_2^*, \ldots x_n^*\}$

  x_i^* can be the same picked randomly from X

- Repeat B times and compute estimates for various Expected values (variance)

  Each time we have an estimate μ_{B_j}

- compute

$$\bar{\mu} = \frac{1}{B}\sum_{j=1}^{B} \mu_{B_j}$$

- standard deviation

$$STD = \sqrt{\frac{1}{B}\sum_{j=1}^{B}\left(\mu_{B_j} - \bar{\mu}\right)^2}$$

Example

- Gaussian with μ and σ^2

  $n = 10000$ numbers

  $X = \{x_1, x_2, \ldots x_{10000}\}$

  $Y = \{y_1, y_2, \ldots y_{10000}\}$

  $\vdots$

- or use Bootstrap m times on $X = \{x_1 x_2 \ldots x_{10000}\}$

Blocking method:

Definitions