* Worksheet 9 :

* problem 1 : $P_r(\text{real}) = 0.9 = p$

$X = X_1 + \cdots + X_n$, $X_i = \begin{cases} 1 & p=0.9 \\ 0 & p=0.1 \end{cases} \Rightarrow \mu = np = 900 \times 0.9 = 810$

$\sigma^2 = np(1-p) = 900 \times 0.9 \times 0.1 = 81 \Rightarrow X \sim N(810, 81)$ Gaussian

problem 3 : $X_i = Z_1 + \cdots + Z_{100}$ $\quad Z_i = \begin{cases} 1 & \text{if in analysis with prob } 1/20 \\ 0 & \text{if not } \cdots \cdots \ 19/20 \end{cases}$

a) $E\{X_i\} = np = 100 \times \frac{1}{20} = 5$

$\text{Var}\{X_i\} = np(1-p) = 100 \times \frac{1}{20} \times \frac{19}{20} = \frac{19}{4}$

b) 95% confidence $\approx 2\sigma \Rightarrow$ upper bound on $X_i = \mu + 2\sigma$

$= 5 + 2\sqrt{\frac{19}{4}}$

$= 5 + \sqrt{19}$

c) $E\{Y_i\} = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$

$\text{Var}\{Y_i\} = \frac{1}{2}(1)^2 + \frac{1}{2}(-1-0)^2 = 1$

d) $\mu = E\{Z_r - Z_b\} = 100 \times 0 = 0$

$\sigma^2 = \text{Var}\{Z_r - Z_b\} = 100 \times 1 = 100$

e) 99% confidence interval for $|Z_r - Z_b| = \mu + 3\sigma = 0 + 3(10) = 30$

due to absolute $\Rightarrow$ can't happen

problem 5 : $p = 0.5$

a) $\text{std} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{\frac{1}{2}\cdot\frac{1}{2}}{100}} = \frac{1}{20}$

b) $\text{std} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{\frac{1}{4}}{2500}} = \frac{1}{100}$

problem 8 : $\hat{p} = \frac{194}{500} = 0.388$

$\bar{X} = \frac{X_1 + \cdots + X_{500}}{500} \Rightarrow E[\bar{X}] = 0.388$

$\bar{X}$ ← percentage of all people

$\text{Var}[\bar{X}] = \frac{p(1-p)}{500} = \frac{0.388 \times 0.612}{500} = 0.00047$

$\Rightarrow$ 95% confidence interval $= 0.388 \pm 2\sqrt{0.00047}$

$= 0.388 \pm 0.0433$

problem 9 : what's important is the characteristics of the demographics $\Rightarrow$ demographics are same $\Rightarrow$ sample size can be same $\Rightarrow n = 1000$

problem 10 : $E[X_i] = 307$ $\quad \text{std}[X_i] = 30$

nation wide average score $= \bar{X} = \frac{X_1 + \cdots + X_{1000}}{1000}$

$\Rightarrow E[\bar{X}] = 307$

$\text{std}[\bar{X}] = \frac{30}{\sqrt{n}} = \frac{30}{\sqrt{1000}} = 0.9488$

Problem 11 : $X = X_1 + \cdots + X_{100}$

a) average numbers in the box $= \frac{X_1 + \cdots + X_{100}}{100} = E[X_i]$

$= \frac{297}{100} = 2.97$

b) $\text{Var}\left[\frac{X_1 + \cdots + X_{100}}{100}\right] = \frac{\text{Var}[X_i]}{100}$ → don't have variance of numbers and sum so we can't give confidence interval