Examination for UE3.3 « Refresher courses »: Statistics Applied to Biology

Isabelle Rivals, Équipe de Statistique Appliquée, ESCPI Paris Tech

Solution

1 Fight against doping

a) We have$P(D) = 0.01$, $P(+|D) = 0.99$, $P(+|D^c) = 0.05$. Applying Bayes' formula :

$$P(D|+) = \frac{P(+|D)P(D)}{P(+|D)P(D) + P(+|D^c)P(D^c)} = \frac{0.99 \times 0.01}{0.90 \times 0.01 + 0.05 \times 0.99} = \frac{1}{6} = 16.7\%$$

b) The conclusion of the paper is wrong : there will be 83.3% of judiciary errors.

2 Cell size statistics

a) Unbiased estimate of the mathematical expectation :

$$\bar{y} = \frac{1}{n}\sum_{k=1}^{n} y^k = 2.63$$

Unbiased estimate of the variance :

$$s^2 = \frac{1}{n-1}\sum_{k=1}^{n}(y^k - \bar{y})^2 = 7.28$$

b) Small sample confidence interval, assuming gaussian measurements :

$$\mu \in \bar{y} \pm t_{n-1}(1-\alpha/2)\frac{s}{\sqrt{n}}$$

A.N. $n = 6$, $\alpha = 0.05$, $t_5(0.975) = 2.57$ (half width 2.83), hence $\mu \in [-0.20; 5.46]$.

c) This interval includes negative values, which are impossible for cell sizes.

d) Confidence interval for the log of the expectation :

$$\mu_z \in \bar{z} \pm t_{n-1}(1-\alpha/2)\frac{s_z}{\sqrt{n}}$$

A.N. $s_z^2 = 0.746 = 0.86^2$ (half width 0.90), hence $\mu_z \in [-0.28; 1.53]$.

e) Correct confidence interval for the log of the size :

$$\mu \in [e^{-0.28}; e^{1.53}]$$

A.N. Hence $\mu \in [0.76; 4.63]$, which does not include 0.