# Assignment 2 – The paper and pencil problems

## Strontium

- Express $h$ in terms of $\lambda$:

  $P(T > h) = \frac{1}{2} \Rightarrow P(T \leq h) = 1 - \frac{1}{2} = \frac{1}{2}$

  $1 - e^{-\lambda h} = \frac{1}{2}$

  $e^{-\lambda h} = \frac{1}{2}$

  $-\lambda h = \ln\frac{1}{2} = \ln 1 - \ln 2 = -\ln 2$

  $\lambda h = \ln 2$

  $\boldsymbol{h = \frac{\ln 2}{\lambda}}$

  The value of $\lambda$ for strontium 90:

  $\boldsymbol{\lambda = \frac{\ln 2}{h} = \frac{\ln 2}{28} \approx 0.025}$

- The probability that a given atom of strontium 90 has not decayed after 50 years:

  Tis Time it remains radioactive > 50

  $\boldsymbol{P(T > 50) = 1 - P(T \leq 50) = 1 - \left(1 - e^{-0.025 \cdot 50}\right) = e^{-1.25} = 0.287}$

- Calculate the number of years to wait so that 99% of the strontium 90 produced during a nuclear reaction disappears:

  We need to find x = the number of years which after the probability of strontium decayed is 0.99:

  $0.99 = 1 - e^{-\lambda x}$

  $e^{-\lambda x} = 0.01$

  $-\lambda x = \ln 0.01 \quad \Rightarrow \quad -\frac{\ln 2}{28} x = \ln 0.01 \quad \Rightarrow \quad \boldsymbol{x = 186}$

## Portfolio

- The expected value of the total amount of money this investor will have after a year:

  $\mathbf{E}(\boldsymbol{X}) = \mathbf{2000 \cdot 110\% + 1000 \cdot 116\% = 2200 + 1160 = 3360\$}$

- The standard deviation of the total amount after a year:

  Var(fixed + fund) = Var(fixed)+Var(fund) assuming that they are independent

  Var(fixed)=0

  **std(total) =std(fund) = 0.08*1000 = 80$**

- Probability that the total wealth will be less than $3000**:**

  For total wealth of 3,000$, the fund value must be: $3000 - 2200 = 800\$$

  Number of standard deviations from the fund expected value in this case:

  $$Z = \frac{800 - 1160}{80} = -4.5$$

  **And the probability is** $\approx \mathbf{0}$

## Plane

- The distribution of $X$ is the **Binomial Distribution:** $\boldsymbol{X \sim B(N, 0.9)}$

- Approximate the probability $P(X \geq n)$:

  The exact distribution is binomial, but it can be approximated with normal precisely enough:

  $\mu = 0.9N$

  $\sigma = \sqrt{N \cdot 0.9 \cdot 0.1} = \sqrt{0.09N} = 0.3\sqrt{N}$

  So, the probability is:

  $$\mathbf{1 - \int_0^n \frac{1}{\sqrt{2\pi \cdot 0.09N}} e^{-\frac{(x-0.9N)^2}{2 \cdot 0.09N}} dx}$$

- Calculate the largest number of bookings the company can accept if it does not want to refuse customers with probability more than 0.05:

  Given P(x<Z) = 0.95,Z = 1.645:

So, to find N = largest number of bookings, we calculate:

$$0.9N + 1.65 \cdot 0.3\sqrt{N} = 380$$

$$\boldsymbol{N \approx 411}$$

- The probability that at least 35 seats are left unoccupied:

$$X \sim B(411{,}0.9) \approx \ \sim N(370, 37)$$

We need to calculate $Pr(X \leq 380 - 35) = Pr(X \leq 345)$.

Using binomial distribution calculator:

$$\boldsymbol{Pr(X \leq 345) = 0.0000892399}$$

Using the Normal distribution:

$$Z = \frac{345 - 370}{\sqrt{37}} = -4.12 \quad \Rightarrow \boldsymbol{P = 0.00002}$$

## Worms

- The unbiased estimators:

**Estimation of true mean:**

$$E\hat{\theta} = \frac{5 + 25 + 15 + 10 + 7 + 12 + 16 + 20}{8} = \mathbf{13.75}$$

**Estimation of variance:**

$$S^2 = \frac{1}{7}\sum (X_i - 13.75)^2$$

$$= \frac{1}{7} \cdot (76.5625 + 126.5625 + 1.5625 + 14.0625 + 45.5625 + 3.0625 + 5.0625 + 39.0625) = \mathbf{44.5}$$

- We can estimate the variance and the mean of our estimate of the mean using the sample distribution of the means. For this distribution we get that

$E(\bar{x})$=μ hence the mean of our estimate of the mean is the same as our estimated mean = $\mathbf{13.75}$

And $Var(\bar{x}) = \frac{\sigma^2}{n}$ but as we don't have $\sigma^2$ - the population variance we can just estimate it using $S^2$ . so we get :

$$\boldsymbol{\sigma_{\bar{x}}^2 = \frac{44.5}{8} = 5.5625}$$

- 95% confidence interval for the true mean:

  The sample is small so we will use t test. the degree of freedom is 7 and we want 95% confidence interval so we need to multiply the standard error by 2.365.

  The confidence interval is:

  $$13.75 - 2.365 \cdot \sqrt{5.5625} \ to \ 13.75 + 2.365 \cdot \sqrt{5.5625}$$

  $$\boldsymbol{\Leftrightarrow 8.173 \ to \ 19.328}$$

## Maximum likelihood with cities

1. 1.1 Derive the PDF by differentiating the CDF with respect to x:

   $$PDF\ f(x) = \left(1 - \left(\frac{x}{a}\right)^{-b}\right)' = \left(1 - \frac{x^{-b}}{a^{-b}}\right)' = a^b b x^{-b-1} = \boldsymbol{\frac{a^b b}{x^{b+1}}} \quad \boldsymbol{[x \geq a]}$$

   1.2The  log-likelihood function:

   $$L(b|x_n) = \prod_{i=1}^{n} \frac{k^b b}{x_i^{b+1}} = \left(b^n k^{nb}\right) \prod_{i=1}^{n} x_i^{-(b+1)}$$

   Log likelihood:

   $$\boldsymbol{\ln L(b|x_n) = n \ln b + nb \ln a - (b+1) \ln\left(\prod_{i=1}^{n} x_i\right)}$$

   $$\boldsymbol{\ln L(b|x_n) = n \ln b + nb \ln a - (b+1) \sum_{i=1}^{n} \ln (x_i)}$$

   1.3The parameters that maximize the log-likelihood:

   - **ln(a) is monotonically increasing. So in order to maximize the likelihood we need to set a as high as possible. Since $\boldsymbol{a \leq x_i}$for all i, we maximize the likelihood by setting ato the smallest $\boldsymbol{x_i}$ in the sample which is $\boldsymbol{\min (x_i)}$.**
   - A value of $b$ at which the derivative of $LL$ with respect to $b$ is 0:

$$\frac{\partial \ell}{\partial \mathrm{b}} = \frac{n}{b} + n \ln a - \ln\left(\sum_{i=1}^{n} x_i\right) = \frac{n}{b} + n \ln a - \sum_{i=1}^{n} \ln\,(x_i)$$

Equals to zero:

$$\frac{\partial \ell}{\partial \mathrm{b}} = 0 \Rightarrow \frac{n}{\hat{b}} + n \ln a - \sum_{i=1}^{n} \ln\,(x_i) = 0$$

$$\frac{n}{\hat{b}} = \sum_{i=1}^{n} \ln\,(x_i) - n \ln a$$

$$\frac{1}{\hat{b}} = \frac{\sum_{i=1}^{n} \ln\,(x_i) - n \ln a}{n}$$

$$\boldsymbol{\hat{b}} = \frac{\boldsymbol{n}}{\boldsymbol{\sum_{i=1}^{n} \ln\,(x_i) - n \ln a}}$$

2. Estimation of the parameters:

   Estimation of $a = \min(x_i) = \mathbf{9018}$

   Estimation of $b =$

$$\frac{\boldsymbol{n}}{\boldsymbol{\sum_{i=1}^{n} \ln\,(x_i) - n \ln a}} = \frac{\mathbf{1001}}{\mathbf{10589.88538 - 1001 \cdot \ln 9018}} = \mathbf{0.679}$$

3. Construct a 99% confidence interval for b:

   For the normal distribution:

   Mean = b

$$LL''_{bb} = \frac{\partial(\frac{n}{b} + n \ln a - \sum_{i=1}^{n} \ln\,(x_i))}{\partial \mathrm{b}} = -nb^{-2} = -\frac{n}{b^2}$$

$$variance = -(LL''_{bb})^{-1} = -\left(-\frac{n}{b^2}\right)^{-1} = -\left(-\frac{b^2}{n}\right) = \frac{b^2}{n}$$

   Approximately: $\hat{b} \sim N(b, \frac{b^2}{n})$

   **99% normal CI for b:** $\boldsymbol{\hat{b}} \pm \mathbf{2.576} \cdot \frac{b}{\sqrt{n}}$

For the cities data:

Mean = 0.679

$$std = \frac{0.679}{\sqrt{1001}} = 0.02146$$

**99% normal CI for b:** $\mathbf{0.679 \pm 2.576 \cdot 0.02146 = 0.679 \pm 0.0553}$