Problem 2: Maximum Likelihood Estimation

1) Given the probability of the data set, given the two parameters.

A: • The likelihood function for a normal distribution is given by the product of the probability density function (PDF) evaluated at each data point. The PDF for a normal distribution is given by:

$$N(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

where: x: The data points

μ: Mean

σ^2: Variance

• Assuming data points are independent and identically distributed from a normal distribution, the probability of the dataset given the two parameters is —

$$P(x|\mu,\sigma^2) = \prod_{n=1}^{10} N(x_n|\mu,\sigma^2) \quad \text{—— (1)}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2\right) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x_2-\mu}{\sigma}\right)^2\right) \cdots$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x_{10}-\mu}{\sigma}\right)^2\right)$$

$$P(x|\mu,\sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{10} \cdot \exp\left(\frac{-1}{2\sigma^2}\sum_{n=1}^{10}(x_n-\mu)^2\right)$$

This is the likelihood function.

2) Derive and calculate the solution for both μ and σ^2 using Maximum likelihood function.

A:- From the previous solution, using equation (1),

$$P(x|\mu,\sigma^2) = \prod_{n=1}^{10} N(x_n|\mu,\sigma^2)$$

$$= \prod_{n=1}^{10} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

Applying log on both sides, for log-likelihood function,

$$\ln(P(x|\mu,\sigma^2)) = \sum_{n=1}^{10} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$= \sum_{n=1}^{10} \cancel{\ln(1)} - \ln(\sqrt{2\pi\sigma^2}) - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$= \left[\sum_{n=1}^{10} \frac{-1}{2}\left(\frac{x_n-\mu}{\sigma}\right)^2\right] - \frac{10}{2}\left[\ln(2\pi) + \ln(\sigma^2)\right]$$

$$= \frac{-1}{2\sigma^2}\sum_{n=1}^{10}(x_n-\mu)^2 - 5\ln(2\pi) - 5\ln(\sigma^2)$$

Now, Substituting the data points for the corresponding log-likelihood functions,

$$\ln(P(x|\mu,\sigma^2)) = \frac{-1}{2\sigma^2}\Big[(112-\mu)^2 + (120-\mu)^2 + (131-\mu)^2 + (126-\mu)^2 + (145-\mu)^2 + (158-\mu)^2 + (157-\mu)^2 + (136-\mu)^2 + (148-\mu)^2 + (176-\mu)^2\Big] - 5\ln(2\pi) - 5\ln(\sigma^2) \quad \text{—— (2)}$$

Taking partial derivative w.r.t μ,

$$\frac{\partial}{\partial\mu}\Big[\ln(P(x|\mu,\sigma^2))\Big] = \frac{-1}{2\sigma^2}\Big[2(112-\mu)(-1) + 2(120-\mu)(-1) + 2(131-\mu)(-1) + 2(126-\mu)(-1) + 2(145-\mu)(-1) + 2(158-\mu)(-1) + 2(157-\mu)(-1) + 2(136-\mu)(-1) + 2(148-\mu)(-1) + 2(176-\mu)(-1)\Big]$$

$$\text{—— (3)}$$

Now, setting the derivative to 0

$$0 = \frac{1}{\sigma^2}\Big[112 + 120 + 131 + 126 + 145 + 158 + 157 + 136 + 148 + 176 - 10\mu\Big].$$

$\Rightarrow \quad 1409 - 10\mu = 0$

$\Rightarrow \quad 10\mu = 1409$

$\therefore \boxed{\mu = 140.9}$ is the mean of the distribution.

Taking partial derivative of equation 2, w.r.t. σ

$$\frac{\partial}{\partial \sigma}\left[\ln\left(P(x|\mu,\sigma^2)\right)\right] = \frac{1}{\sigma^3}\Big[(112-140.9)^2+(120-140.9)^2+(131-140.9)^2$$

$$+(126-140.9)^2+(145-140.9)^2+(158-140.9)^2+(157-140.9)^2$$

$$+(136-140.9)^2+(148-140.9)^2+(176-140.9)^2\Big] - 5\times\frac{1}{\cancel{\sigma^2}}\times 2\cancel{\sigma}$$

Setting the derivative to 0,

$$0 = \frac{1}{\sigma^3}\Big[835.21+436.81+98.01+222.01+16.81+292.41+259.21+24.01+50.41+1232.01\Big] - \frac{10}{\sigma}$$

$$\frac{10}{\cancel{\sigma}} = \frac{1}{\sigma^{\cancel{3}\,2}}(3466.9)$$

$\boxed{\sigma^2 = 346.69}$ is the variance of the distribution.