Ex 1-2: See PN.

Ex 3:

The ML estimate is

$$\hat{x} = \arg\max_x f_{Y|X}(y|x).$$

Here, we have

$$f_{Y|X}(y|x) = x(1-x)^{y-1}, \quad y = 1, 2, \ldots$$

So we want to maximize (note $Y=5$)

$$x(1-x)^{y-1} = x(1-x)^4$$

$$\frac{d}{dx} x(1-x)^4 = 5(1-x)^4 - 4(1-x)^3 = 0$$

$$\Leftrightarrow 5(1-x)^4 = 4(1-x)^3$$

$$\Leftrightarrow 5(1-x) = 4$$

$$\Leftrightarrow \hat{x}_{ML} = \frac{1}{5}$$

For the MAP estimator, we have

$$f_{X|Y}(x|y) = f_{Y|X}(y|x) f_X(x) = x(1-x)^4 3x^2$$

$$= 3x^3(1-x)^4$$

Differentiating to optimize, we get

$$\frac{d}{dx} f_{X|Y}(x|y=5) = 3x^2(1-x)^4 - 4(1-x)^3 x^3 = 0$$

$$\Leftrightarrow \hat{x}_{MAP} = \frac{3}{7}$$

hence, the MAP is quite a bit higher than ML in this case, which is reasonable since the prior favors larger values of x.

Ex 4 :

Using the hint, we have

$$E[X|Y] = \frac{\sigma_X}{\sigma_Y} \rho_{XY} Y$$

Fact: For $Z \sim N(\mu, \sigma^2)$

$$aZ \sim N(a\mu, a^2\sigma^2)$$

$$\Rightarrow E[X|Y] \sim N\left(0, \frac{\sigma_X^2}{\sigma_Y^2} \rho_{XY}^2 \sigma_Y^2\right)$$

$$= N\left(0, \sigma_X^2 \rho_{XY}^2\right)$$

$$\hat{X}_{LMMSE} = \frac{cov(X,Y)}{var(Y)}(Y-\mu_Y)+\mu_X$$

$$= \frac{\sigma_{XY}}{\sigma_Y^2} Y$$

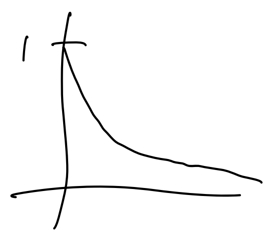
$$= \frac{\rho_{XY}\sigma_X\sigma_Y}{\sigma_Y^2} Y = \frac{\sigma_X}{\sigma_Y}\rho_{XY} Y$$

$$\hat{X}_{MAP} = \arg\max_x f_{X|Y}(x|y)$$

$$= \arg\max_x \mathcal{N}\left(\frac{\sigma_X}{\sigma_Y}\rho_{XY} y,\ \overbrace{\sigma_X^2(1-\rho_{XY})}^{\sigma^2}\right)$$

$$= \arg\max_x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left(x-\frac{\sigma_X}{\sigma_Y}\rho_{XY} y\right)^2\right)$$

$$= \frac{\sigma_X}{\sigma_Y}\rho_{XY} y \qquad e^{-k},\ k \geq 0$$



Ex 5 :

Since Y_i's are iid, have

$$f_{Y_1 \dots Y_N}(y_1, \dots, y_N) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}$$

Likelihood function

$$f_{Y_1 \dots Y_N}(y_1 \dots y_N ; \mu, \sigma^2) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}$$

Log likelihood : $\log(ab) = \log(a) + \log(b)$

$$\ell(\mu, \sigma^2) = \sum_{i=1}^{N} \left(\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{(y_i-\mu)^2}{2\sigma^2} \right)$$

$$= -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{N} (y_i - \mu)^2$$

$$\frac{d}{d\mu} \ell(\mu, \sigma^2) = -\cancel{2} \frac{1}{\cancel{2}\sigma^2} \sum_{i=1}^{N} (y_i - \mu) = 0$$

$$\Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^{N} y_i = \frac{1}{\sigma^2} \sum_{i=1}^{N} \mu \Rightarrow \sum_{i=1}^{N} y_i = N\mu$$

$$\Leftrightarrow \boxed{\hat{\mu}_{ML} = \frac{1}{N} \sum_{i=1}^{N} y_i}$$

Now use $\hat{\mu}$ to find $\hat{\sigma}$:

$$\frac{d}{d\sigma^2} \ell(\mu, \sigma^2) = -\frac{N}{2} \frac{2\pi}{2\pi\sigma^2} + \frac{1}{\sigma^4} \sum_{i=1}^{N} \frac{(y_i-\mu)^2}{2} = 0$$

$$\Rightarrow N = \frac{1}{\sigma^2} \sum_{i=1}^{N} (y_i-\mu)^2$$

$$\Rightarrow \hat{\sigma}^2_{ML} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{\mu}_{ML})^2$$

$$\mathrm{Var}(Y) = \mathbb{E}\left[(Y - \mathbb{E}[Y])^2\right]$$

Ex 6:

Since the Y_i's are iid, their joint PDF is the product of the individual $f_Y(y) = \frac{\lambda^y}{y!}e^{-\lambda}$. Further, since λ is fixed,

$$f_Y(y;\lambda) = \frac{\lambda^y}{y!}e^{-\lambda}$$

denotes that this isn't really a conditional distribution

and therefore

$$f_{Y_1,\dots,Y_N}(y_1,\dots,y_N;\lambda) = \prod_{i=1}^{N} \frac{\lambda^{y_i}}{y_i!}e^{-\lambda}$$

We want to optimize this over λ, so it will be easier to take the log first. We call this the <u>log-likelihood</u> $\ell(\lambda)$

$$\ell(\lambda) = \log\left(\prod_{i=1}^{N} \frac{\lambda^{y_i}}{y_i!}e^{-\lambda}\right)$$

$$= \sum_{i=1}^{N} \log\left(\frac{\lambda^{y_i}}{y_i!}e^{-\lambda}\right)$$

$$= \sum_{i=1}^{N} y_i\log(\lambda) - \log(y_i!) - \lambda$$

Note that $y_i!$ doesn't depend on λ, so we can ignore it.

$$\hat{\lambda} = \underset{\lambda}{\operatorname{argmax}} \sum_{i=1}^{N} y_i \log(\lambda) - \lambda$$

$$= \underset{\lambda}{\operatorname{argmax}} \log(\lambda)\left(\sum_{i=1}^{N} y_i\right) - N\lambda$$

Now differentiate and set to zero

$$\frac{\partial}{\partial \lambda} \ell(\lambda) = \frac{1}{\lambda} \sum_{i=1}^{N} y_i - N$$

$$\Rightarrow \quad \hat{\lambda} = \frac{1}{N} \sum_{i=1}^{N} y_i$$

This is the sample mean, which makes sense because

$$\mathbb{E}[Y_i] = \lambda$$

for all i.