def sample n times.

Prob 3. $\hat{y}_{ml} = \arg\max_y p(x|y) = \arg\max_y \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp(\frac{(x_i-y)^2}{2})$

(a)

$$= \arg\max_y \ln\left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp(\frac{(x_i-y)^2}{2})\right) = \arg\max_y \left(\sum_{i=1}^{n} -\frac{(x_i-y)^2}{2}\right)$$

$$= \frac{\sum_{i=1}^{n} x_i}{n}$$

for MAP, we maximize $\ln(p(x|y)\,p(y))$. i.e.

$$\hat{y}_1 = \arg\max_y \ln[p(x|y)\,p(y)] =$$

$$\arg\max_y \left\{A + \sum_{i=1}^{n}\left(-\frac{(x_i-y)^2}{2}\right) + \left(-\frac{s(y-z)^2}{2}\right)\right\} \triangleq f \quad A = \text{CONST}$$

$\frac{\partial f}{\partial y} = 0$, we have $\Rightarrow \sum_{i=1}^{n}(x_i - y) + s(y-z) = 0$

$$\hat{y}_1 = \frac{\sum_{i=1}^{n} x_i + sz}{n+s} = \frac{\hat{y}_{ml}\cdot n + sz}{n+s}$$

(b). prior distro of y is Uniform, $p(x|y)\,p(y)$, $p(x,y)$

in the domain, $y \in [-1,1]$.

$$\arg\max_{y\in D} p(x|y)\,p(y) = \arg\max_{y\in D} p(x|y)$$

indicating in the Uniform domain $[-1,1]$ MLE is

same as MAP-Estimator.

$$\hat{y}_2 = \begin{cases} \hat{y}_{MLE} = \frac{\sum_{i=1}^{n} x_i}{n} & (\sum_{i=1}^{n} x_i \in [-n, n] \\ \text{not defined} & (\sum_{i=1}^{n} x_i \notin [-n, n] \end{cases}$$

(c) Real y distro $y \sim N(0,2)$

+choose $\hat{y}_1|_{z=10,s=1}$ or $\hat{y}_2$ for $n \to 1$ / $n \to \infty$

$n \to 1$,

$\hat{y}_2 = \hat{y}_{MLE}$ so long as $\frac{\sum_{i=1}^{1} x_1}{1}$ is in the domain,

which is probably guaranteed, while $\hat{y}_1|_{z=10,s=1} = \frac{\hat{y}_{MLE}+10}{2}$

Hence $\hat{y}_2$ is more like $\hat{y}_{MLE}$ and should be chosen.

$n \to \infty$ both are close to $\hat{y}_{MLE}$, while $\hat{y}_1$ estimator

has a wider domain, rather than $[-1,1]$, chosen.

Prob 4. def $A = \int_{\omega_{Dom}} L(y,\omega)\, p(\omega|x)\, d\omega$

Minimum loss (conditional) is obtained when $\frac{\partial A}{\partial y} = 0$

$$\frac{\partial A}{\partial y} = \frac{\partial}{\partial y}\left[\int_{\omega_{Domain}} (y-\omega)^2 p(\omega|x)\, d\omega\right] = \int_{\omega_{Domain}} \frac{\partial[(y-\omega)^2]}{\partial y}\, p(\omega|x)\, d\omega$$

$$= \int_{\omega_{Domain}} 2(y-\omega)\, p(\omega|x)\, d\omega = 0$$

$$\Rightarrow y\Big(\underbrace{\int_{\omega_{Domain}} p(\omega|x)\, d\omega}\Big) = \underbrace{\int_{\omega_{Domain}} \omega\, p(\omega|x)\, d\omega}$$

$$\boxed{y \cdot 1 = E(\omega|x)}$$

Estimator which brings "min-risk(conditional)"

is the conditional expectation of the A-Post-distro

Overall risk is minimized so long as conditional risk is minimized. Thus the two estimators, which bring either min-overall-risk or min-conditional-risk, should be identical.

b. same procedure.

$$0 = \frac{\partial A}{\partial y} = \frac{\partial}{\partial y}\left[\int_{\omega_D} |y-\omega|\, p(\omega|x)\, d\omega\right] = \frac{\partial}{\partial y}\Big[\int_{\omega>y} (\omega-y)\, p(\omega|x)\, d\omega + \int_{\omega<y} (y-\omega)\, p(\omega|x)\, d\omega\Big]$$

$$0 = -\int_{\omega>y} p(\omega|x)\, d\omega + \int_{\omega<y} p(\omega|x)\, d\omega \Leftrightarrow \int_{\omega<y} p(\omega|x)\, dx = \int_{\omega>y} p(\omega|x)\, dx$$

$\Downarrow$

y is median

since for "$\omega<y$"

"$\omega>y$"

share equal prob.

ii. same mentality as in (a)-session.

Prob 5.

a.

$$\hat{\lambda}_{MLE} = \arg\max_{\lambda} \left(\prod_{i=1}^{N} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \right)$$

$$= \arg\max_{\lambda} \left(\ln \prod_{i=1}^{N} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \right)$$

$$= \arg\max_{\lambda} \sum_{i=1}^{N} (x_i \log\lambda - \lambda)$$

$$\frac{\partial \left(\sum_{i=1}^{N} x_i \log\lambda - \lambda \right)}{\partial \lambda} = \sum_{i=1}^{N} \left(x_i \frac{1}{\lambda} - 1 \right) \overset{def}{=} 0 \Rightarrow \hat{\lambda}_{MLE} = \frac{\sum_{i=1}^{N} x_i}{N}$$

$$E\left[\sum_{i=1}^{N} x_i / N\right] = \frac{1}{N} \sum_{i=1}^{N} E[x_i] = \frac{N\lambda}{N} = \lambda \Rightarrow \text{unbiased.}$$

---

b. $p = P(X=0)^2 = e^{-2\lambda}$ thus $\lambda = -\frac{1}{2}\ln p$

single sample, thus no "$\prod$" i.e.

$$\hat{p} = \arg\max_{p} p(x|p) = \arg\max_{p} \underbrace{\ln \left[\frac{(-\frac{1}{2}\ln p)^x}{x!} e^{\frac{1}{2}\ln p} \right]}_{\text{"}A\text{"}}$$

$$\frac{\partial A}{\partial p} = 0 \text{ i.e. } \frac{1}{2} + \frac{x}{\ln p} = 0 \therefore \hat{p} = e^{-2x}$$

$$E(\hat{p}) = E(e^{-2x}) = \sum_{x_i=0}^{+\infty} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \cdot e^{-2x_i}$$

Given $\left\langle e^u = 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!} + \cdots + \frac{u^n}{n!} + \cdots \right\rangle$

$$\sum_{x_i=0}^{+\infty} \frac{(\lambda e^{-2})^{x_i}}{x_i!} = e^{\lambda e^{-2}}$$

$$\therefore E(\hat{p}) = e^{-\lambda + \lambda e^{-2}} \neq e^{-2\lambda} \Rightarrow \text{biased}$$

---

c.