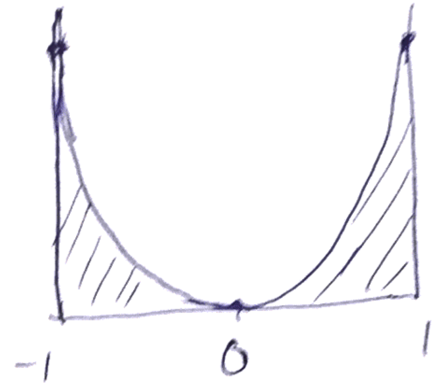
Solution Problem Set 2

1. [redacted]



[redacted]

$$p(y \mid x) = \frac{1}{x^2} \quad 0 \le y \le x^2$$
$$= 0$$

$$p(x) = \frac{1}{2}$$

$$p(x,y) = \frac{1}{2x^2} \quad \text{for } \left.\begin{array}{l} -1 \le X \le 1 \\ 0 \le y \le x^2 \end{array}\right\} \text{Region } R$$

a) $E[X] = 0$ given symmetry

$(x, y) \to (-x, y)$ is a symmetry of the distr$\underline{n}$.

$$E[X\,Y] = \int_R dx\,dy\; xy\, p(x,y)$$
$$= 0$$

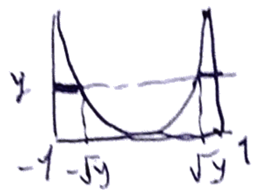
Using $x \to -x$ transformation

So $\mathrm{Cov}(X,Y) = E[X\,Y] - E[X]E[Y] = 0$

$\Rightarrow$ correlation $= 0$

b) $I(x,y) = H[y] - H(y|x)$

$$p(y) = \left[\int_{-1}^{-\sqrt{y}} + \int_{\sqrt{y}}^{1}\right] \left[\frac{1}{2x^2}\right] dx$$



$$= 2\left[-\frac{1}{2x}\right]_{\sqrt{y}}^{1} = \frac{1}{\sqrt{y}} - 1 = \frac{1-\sqrt{y}}{\sqrt{y}}$$

$$H(Y) = \int_0^1 dy \, \frac{1-\sqrt{y}}{\sqrt{y}} \log \frac{\sqrt{y}}{1-\sqrt{y}} = 2\int_0^1 dt \, (1-t) \log \frac{t}{1-t}$$

with $y = t^2$

$$H(Y|X) = \frac{1}{2}\int_{-1}^{1} dx \, \underbrace{\log x^2}_{\text{entropy of } y \text{ for fixed } x} = 2\int_0^1 dx \log x$$

So
$$H(Y) - H(Y|X) = -2\int_0^1 dt \, (1-t)\log(1-t) + 2\int_0^1 dt \, (1-t)\log t - 2\int_0^1 dx \log x$$

$$= -2\int_0^1 dt \, (1-t)\log(1-t) + \cancel{2\int_0^1 dt \log t} - 2\int_0^1 dt \, t \log t \; \cancel{- 2\int_0^1 dx \log x}$$

$$= -4\int_0^1 dt \, t \log t$$

$$\int_0^1 t^{1+\alpha}\,dt = \left[\frac{t^{2+\alpha}}{2+\alpha}\right]_0^1 = \frac{1}{2+\alpha} = \frac{1}{2}\left(1+\frac{\alpha}{2}\right)^{-1} = \frac{1}{2} - \frac{\alpha}{4} + o(\alpha^2)$$

$$\int_0^1 t^{1+\alpha}\,dt = \int_0^1 t\,(1+\alpha\log t + o(\alpha^2))\,dt$$

$$= \frac{1}{2} + \alpha\int_0^1 dt\; t\log t + o(\alpha^2)$$

So $\int_0^1 dt\; t\log t = -\frac{1}{4}$

$I(X,Y) = 1$ in nats (since I use log base e)

2. All of you did a fine job generating the distribution of the estimates and getting it to be peaked close to the true values.

3.a) $$P(x|\sigma^2) = \frac{e^{-\frac{\sum_{i=1}^{N} x_i^2}{2\sigma^2}}}{(2\pi\sigma^2)^{N/2}}$$

$$\log P = -\frac{\sum_{i=1}^{N} x_i^2}{2\sigma^2} - \frac{N}{2}\log 2\pi\sigma^2$$

Vary w.r.t. σ^2

$$\frac{\sum_i x_i^2}{2(\sigma^2)^2} - \frac{N}{2\sigma^2} = 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{N} x_i^2}{N}$$

$$E[\hat{\sigma}^2] = \frac{N\sigma^2}{N} = \sigma^2$$

b) $$\mathrm{Var}[\hat{\sigma}^2] = \mathrm{Var}\left[\frac{\sum_i (x_i^2 - E[x_i^2])}{N}\right]$$

since $x_i^2 - E[x_i^2]$ are independent

$$\mathrm{Var}[\hat{\sigma}^2] = \sum_{i=1}^{N} \mathrm{Var}\left[\frac{x_i^2 - E[x_i^2]}{N}\right]$$

$$= \frac{1}{N^2}\sum_i E[(x_i^2 - \sigma^2)^2] = \frac{1}{N^2}\sum_i \left(E[x_i^4] - \sigma^4\right)$$

$$= \frac{1}{N^2} N (3\sigma^4 - \sigma^4) = \frac{2\sigma^4}{N}$$

c).

$$I(\sigma^2) = -E\left[\frac{\partial^2 \log P}{(\partial \sigma^2)^2}\right]$$

$$= -E\left[\frac{\partial^2}{(\partial\sigma^2)^2}\left\{-\frac{\sum x_i^2}{2\sigma^2} - \frac{N}{2}\log(2\pi\sigma^2)\right\}\right]$$

$$= -E\left[-\frac{\sum_i x_i^2}{(\sigma^2)^6} + \frac{N}{2}\frac{1}{(\sigma^2)^2}\right]$$

$$= \frac{N\sigma^2}{(\sigma^2)^6} - \frac{N}{2(\sigma^2)^2}$$

$$= \frac{N}{2\sigma^4}$$

Cramer-Rao bound

$$\mathrm{Var}(\hat{\sigma}^2) \geq \frac{2\sigma^4}{N}$$

In this case we have an equality

4.

$$P(D|\theta) = \prod_{i=1}^{N}\prod_{\alpha=1}^{K}\frac{\theta_\alpha^{x_\alpha^{(i)}}}{x_\alpha^{(i)}!}$$

depend only on data and not θ

$$\log P = \sum_\alpha\left(\sum_i x_\alpha^{(i)}\right)\log\theta_\alpha - \mathrm{Cst}$$

We have the constraint $\sum_{\alpha} \theta_\alpha = 1$

$$\frac{\partial}{\partial \theta_\alpha}\left[\log P(\mathcal{D}|\theta) - \lambda\left(\sum_{\alpha'} \theta_{\alpha'} - 1\right)\right]$$

$$\frac{\sum_i X_\alpha^{(i)}}{\hat{\theta}_\alpha} - \lambda = 0$$

So $\hat{\theta}_\alpha = c \sum_{i=1}^{N} X_\alpha^{(i)}$

To determine c

$\sum_\alpha \hat{\theta}_\alpha = 1$

$$\begin{aligned} 1 &= c \sum_{\alpha=1}^{K}\sum_{i=1}^{N} X_\alpha^{(i)} \\ &= c \sum_{i=1}^{N}\sum_{\alpha=1}^{K} X_\alpha^{(i)} \\ &= c \sum_{i=1}^{N} n \\ &= cNn \end{aligned}$$

So $\hat{\theta}_\alpha = \dfrac{\sum_{i=1}^{N} X_\alpha^{(i)}}{nN}$