Solutions to Homework 10

Exercise 1. We use the large deviations principle to find a tight upper bound. Before this, we need to check that the moment generating function $\mathbb{E}(e^{sX_1})$ is finite in a proper neighborhood of $s = 0$:

$$\mathbb{E}(e^{sX_1}) = \int_0^\infty e^{sx}\,\lambda e^{-\lambda x}\,dx = \frac{\lambda}{\lambda - s}, \quad \text{for } s < \lambda$$

Therefore, by applying the large deviations principle, we obtain for $t > 1/\lambda$:

$$\mathbb{P}(\{S_n > nt\}) \le \exp(-n\,\Lambda^*(t)) \quad \text{where} \quad \Lambda^*(t) = \max_{s\in\mathbb{R}} \left\{ st - \log\left(\frac{\lambda}{\lambda - s}\right)\right\}$$

By taking the derivative of $st - \log\left(\frac{\lambda}{\lambda-s}\right)$ with respect to s and setting it equal to zero, we obtain that $\Lambda^*(t)$ is maximum at $s^* = \lambda - \frac{1}{t}$. Hence,

$$\mathbb{P}(\{S_n > nt\}) \le \exp(-n\,(\lambda t - 1 - \log(\lambda t)))$$

Exercise 2*. a) Use part (ii) of the definition with $U \equiv 1$ (such a U belongs to G).

b) (i) $Z = \mathbb{E}(X)$ is constant and therefore $\mathcal{G}$-measurable; (ii) Let $U \in G$: $\mathbb{E}(XU) = \mathbb{E}(X)\,\mathbb{E}(U) = \mathbb{E}(\mathbb{E}(X)\,U) = \mathbb{E}(ZU)$ (using the independence of X and U and the linearity of expectation).

c) (i) $Z = X$ is $\mathcal{G}$-measurable by assumption; (ii) Let $U \in G$: $\mathbb{E}(XU) = \mathbb{E}(ZU)$!

d) (i) $Z = \mathbb{E}(X|\mathcal{G})\,Y$ is $\mathcal{G}$-measurable; (ii) Let $U \in G$: $\mathbb{E}(XYU) = \mathbb{E}(\mathbb{E}(X|\mathcal{G})\,YU)$, because part (ii) of the definition of $\mathbb{E}(X|\mathcal{G})$ implies the previous equality (indeed, $YU \in G$). Therefore, $\mathbb{E}(XYU) = \mathbb{E}(ZU)$.

e) Let us first check the left-hand side equality: $\mathbb{E}(X|\mathcal{H})$ is $\mathcal{H}$-measurable, therefore $\mathcal{G}$-measurable, so one can apply property c).

For the right-hand side equality, one has: (i) $Z = \mathbb{E}(X|\mathcal{H})$ is $\mathcal{H}$-measurable; (ii) Let $U \in H$:

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})\,U) = \mathbb{E}(\mathbb{E}(XU|\mathcal{G})) = \mathbb{E}(XU) = \mathbb{E}(\mathbb{E}(X|\mathcal{H})\,U) = \mathbb{E}(ZU)$$

using successively d), a) and the definition of $\mathbb{E}(X|\mathcal{H})$.

Exercise 3. a) We must check that $\mathbb{E}(\psi(Y)\,g(Y)) = \mathbb{E}(X\,g(Y))$ for any continuous and bounded function g. The computation gives indeed:

$$\mathbb{E}(\psi(Y)\,g(Y)) = \sum_{y\in C} \psi(y)\,g(y)\,\mathbb{P}(\{Y = y\}) = \sum_{x,y\in C} x\,g(y)\,\mathbb{P}(\{X = x, Y = y\}) = \mathbb{E}(X\,g(Y))$$

b) Let Y and Z be the two independent dice rolls: $\mathbb{P}(\{Y = i\}) = \mathbb{P}(\{Z = j\}) = 0.25$ and $\mathbb{P}(\{Y = i, Z = j\}) = \mathbb{P}(\{Y = i\})\,\mathbb{P}(\{Z = j\})$. We therefore have $\mathbb{E}(\max(Y,Z)|Y) = \psi(Y)$, where

$$\begin{aligned}
\psi(i) &= \sum_{j=i}^4 \max(i,j)\,\mathbb{P}(\{\max(Y,Z) = j\}|\{Y = i\}) = \sum_{j=i}^4 \max(i,j)\,\frac{\mathbb{P}(\{\max(Y,Z) = j, Y = i\})}{\mathbb{P}(\{Y = i\})} \\
&= i\,\frac{\mathbb{P}(\{Z \le i, Y = i\})}{\mathbb{P}(\{Y = i\})} + \sum_{j=i+1}^4 j\,\frac{\mathbb{P}(\{Z = j, Y = i\})}{\mathbb{P}(\{Y = i\})} = i\,\mathbb{P}(\{Z \le i\}) + \sum_{j=i+1}^4 j\,\mathbb{P}(\{Z = j\})
\end{aligned}$$

So $\psi(1) = 2.5$, $\psi(2) = 2.75$, $\psi(3) = 3.25$ and $\psi(4) = 4$.

Exercise 4. a) Let us compute theoretically the first three MSE's:

$$\mathbb{E}((\widehat{X}_1 - X)^2) = \mathbb{E}\left(\left(X + \frac{Z}{a} - X\right)^2\right) = \frac{\mathbb{E}(Z^2)}{a^2} = \frac{1}{a^2}$$

$$\mathbb{E}((\widehat{X}_2 - X)^2) = \mathbb{E}\left(\left(\frac{a^2X + aZ}{a^2+1} - X\right)^2\right) = \mathbb{E}\left(\left(-\frac{1}{a^2+1}\,X + \frac{a}{a^2+1}\,Z\right)^2\right)$$

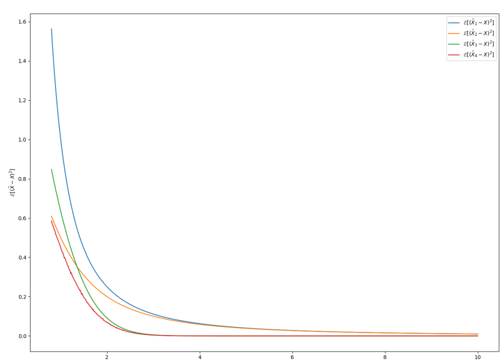
$$= \frac{1}{(a^2+1)^2} + \frac{a^2}{(a^2+1)^2} = \frac{1}{a^2+1}$$

$$\mathbb{E}((\widehat{X}_3 - X)^2) = \mathbb{E}((\text{sign}(a^2X + aZ) - X)^2) = 4Q(-|a|)$$

where $Q(x) = \int_{-\infty}^{x} p_Z(z)dz$ is the cdf of $Z \sim \mathcal{N}(0,1)$. The fourth MSE is given by

$$\mathbb{E}((\widehat{X}_4 - X)^2) = \mathbb{E}((\tanh(a^2X + aZ) - X)^2) = 1 - \mathbb{E}(\tanh(a^2 + aZ)^2)$$

(see computation below in part c) for the last equality). This expression can be computed by numerical integration or by the Monte-Carlo method, as suggested in the problem set. The four MSE's are represented as functions of a on the figure below:



This shows that the fourth estimator gives the minimum MSE.

b) As we shall see below, the fourth estimator corresponds to the conditional expectation $\mathbb{E}(X|Y)$, which by definition minimizes $\mathbb{E}((Z-X)^2)$ among all random variables Z which are $\sigma(Y)$-measurable and square-integrable. The computation of the condition expectation gives $\mathbb{E}(X|Y) = \psi(Y)$, where

$$\psi(y) = \sum_{x\in\{-1,+1\}} x\,\frac{p_x\,p_Z(y-ax)}{p_Y(y)} = \frac{p_Z(y-a) - p_Z(y+a)}{p_Z(y-a) + p_Z(y+a)} = \frac{e^{ay} - e^{-ay}}{e^{ay} + e^{-ay}} = \tanh(ay)$$

which confirms that $\mathbb{E}(X|Y) = \tanh(aY) = \widehat{X}_4$.

NB: The first expression for the function $\psi(y)$ above can be obtained either by reasoning intuitively (and forgetting that we are dealing here with a mix of discrete (X) and continuous (Y) random variables), or by proving formally that the random variable $\psi(Y)$ satisfies (similarly to Ex. 2.a):

$$\mathbb{E}(X g(Y)) = \mathbb{E}(\psi(Y)\,g(Y)), \quad \text{for every } g:\mathbb{R}\to\mathbb{R} \text{ continuous and bounded}$$

c) Using the following series of equalities:

$$\begin{aligned}\mathbb{E}((\mathbb{E}(X|Y)-X)^2) &= \mathbb{E}(X^2)+\mathbb{E}(\mathbb{E}(X|Y)^2)-2\mathbb{E}(X\mathbb{E}(X|Y))\\ &= \mathbb{E}(X^2)+\mathbb{E}(\mathbb{E}(X|Y)^2)-2\mathbb{E}(\mathbb{E}(X\mathbb{E}(X|Y)|Y))\\ &= \mathbb{E}(X^2)+\mathbb{E}(\mathbb{E}(X|Y)^2)-2\mathbb{E}(\mathbb{E}(X|Y)^2)=\mathbb{E}(X^2)-\mathbb{E}(\mathbb{E}(X|Y)^2)\end{aligned}$$

we see that

$$\mathbb{E}((\widehat{X}_4-X)^2)=\mathbb{E}(X^2)-\mathbb{E}(\widehat{X}_4^2)=1-\mathbb{E}(\tanh(aY)^2)=1-\mathbb{E}(\tanh(a^2+aZ)^2)$$

(noticing for the last equality that the value of X can be replaced by $+1$ using symmetry). A direct computation shows that it also holds that $\mathbb{E}((\widehat{X}_2-X)^2)=\mathbb{E}(X^2)-\mathbb{E}(\widehat{X}_2^2)=\frac{1}{a^2+1}$, but that the equality does not hold for $\widehat{X}_1$ and $\widehat{X}_3$.