

# IT for Statistics and Learning 2025

## Assignment 6

Assigned: Thursday, Dec 11, 2025

Due: Thursday, Dec 18, 2025

M. Skoglund

---

**Problem 6.1:** Prove that if  $P_{S|Z^n}$  is  $\varepsilon$  stable, then it is  $\varepsilon$  information stable for any  $Q$ .

**Problem 6.2:** Prove that if  $\Pr(X \in [a, b]) = 1$  for  $-\infty < a \leq b < \infty$  then  $X$  is  $\sigma^2$ -sub-Gaussian with  $\sigma^2 = (b - a)^2/4$

**Problem 6.3:** Let  $g(S, Z^n) = L_Q(S) - L_{Z^n}(S)$ . Assume that  $\ell(s, Z)$  is  $\sigma^2$ -sub-Gaussian for each  $s$  and let

$$i(S, Z^n) = \ln \frac{dP_{S, Z^n}}{d(P_S \otimes P_{Z^n})}$$

(assuming  $P_{S, Z^n} \ll P_S \otimes P_{Z^n}$ ). Prove that

$$E_{P_{S, Z^n}} \left[ \exp \left( \lambda g(S, Z^n) - \frac{\lambda^2 \sigma^2}{2n} - i(S, Z^n) \right) \right] \leq 1$$

using the change-of-measure formula

$$E_{P_{X, Y}} [f(X, Y)] = E_{P_X \otimes P_Y} [e^{i(X, Y)} f(X, Y)]$$

**Problem 6.4:** Use the bound from Problem 6.3 to prove that

$$|E[g(S, Z^n)]| \leq \sqrt{\frac{2\sigma^2}{n} I(S; Z^n)}$$

if  $\ell(s, Z)$  is  $\sigma^2$ -sub-Gaussian for each  $s$

**Problem 6.5:** Consider the binary classifier for samples in  $\mathbb{R}^2$  that separates points by a straight line, partitioning  $\mathbb{R}^2$  into two half-spaces and assigning one label to one half-space and another label to points in the other half-space. Prove that the VC dimension of this set of classifiers is  $d = 3$ .

**Problem 6.6:** For any  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , let  $F(s, z^n) = f(L_{z^n}(s), L_Q(s))$ ,  $P_{z^n} = P_{S|Z^n=z^n}$  and take  $R$  to be any distribution on  $(E, \mathcal{E})$ . Then from the Gibbs variational principle (Lec. 7)

$$\ln \int e^{nF(s, z^n)} R(ds) \geq n \int F(s, z^n) P_{z^n}(ds) - D(P_{z^n} \| R)$$

for each  $z^n$ . Also, for any  $\beta > 0$  and using Markov's inequality,

$$Q^{\otimes n} \left\{ \int e^{nF(s, Z^n)} R(ds) \geq \frac{1}{\beta} \int \left[ \int e^{nF(s, z^n)} R(ds) \right] Q^{\otimes n}(dz^n) \right\} \leq \beta$$

Hence, with probability at least  $1 - \beta$ ,

$$n \int F(s, Z^n) P_{Z^n}(ds) - D(P_{Z^n} \| R) \leq \ln \frac{1}{\beta} + \ln \int \left[ \int e^{nF(s, z^n)} R(ds) \right] Q^{\otimes n}(dz^n)$$

for all possible  $P_{Z^n} = P_{S|Z^n}$ .

Use this result to show that if we take  $\ell(s, Z)$  to be  $\sigma^2$ -subgaussian for each  $s$ , and with

$$f(x, y) = \frac{n-1}{2\sigma^2 n} (y - x)^2$$

it holds that

$$E[L_Q(S)|Z^n] \leq E[L_{Z^n}(S)|Z^n] + \sqrt{\frac{2\sigma^2}{n-1} \left( D(P_{Z^n} \| R) + \ln \frac{\sqrt{n}}{\beta} \right)}$$

with probability at least  $1 - \beta$ . Thus proving the generic PAC Bayes bound considered in class.