RATIONAL INATTENTION EXERCISE

A price-setter has a loss function $E[(P-C)^2] - \lambda\mathcal{I}$ where C is input cost, P is price, and $\mathcal{I}$ is mutual information between the price-setter's observations X (which he can use to set P as $P(X)$) and C.

(1) Suppose C takes on just two possible values, 4 and 5, and it is known to take them on with equal probability. If the price-setter observes a signal that also takes on just two values, then we can think of the information flow directly as mutual information between the P and C random variables. It is natural to guess that because the probabilities of the two costs are equal and the loss function is symmetric, the solution will make P take on a high and low value (P_h and P_ℓ) that each have the same probability, and that $Pr[P = P_h \mid C = 5] = \alpha = Pr[P = P_\ell \mid C = 4]$. Plot α, expected losses, and $P_h - P_\ell$ as functions of λ. [Analytic expressions for these objects as functions of λ can be found.]

(2) Suppose $C = \sum_i Z_i$, with i running from one to four and that $Z_i \sim N(4, i)$, i.i.d. across i. The "water-filling" theorem in information theory asserts that, if we are going to collect observations in the form $X_i = Z_i + \varepsilon_i$, with ε_i independent of Z_j for all i, j and jointly normal with the Z's, for large λ it will be optimal to collect information only on the X_i's with the largest variance, so that the conditional variance of $Z_i \mid X_i$ is the same for all i for which any observation is collected.

However in the model we are considering, it is not optimal to collect separate observations on the separate Z_i's. It will be optimal to observe $X = \sum_i Z_i + \varepsilon$, where ε is jointly normal with the Z's and independent of them. Show that this approach is better. You can do it either by showing it analytically in general, or by considering the case where with separate observations $\mathrm{Var}(Z_i \mid X_i) = 2$ for $i \geq 2$ and Z_1 is unobserved. Calculate the mutual information between C and P in that case and show that lower losses can be obtained at the same or lower information flow by the $X = \sum Z_i + \varepsilon$ approach.