Review Chapters 8 and 10 from Cover and Thomas before starting the homework.

1. Problem 8.1 from Cover and Thomas: This is an exercise in computing differential entropy.

2. Problem 8.8 from Cover and Thomas: This demonstrates computation of mutual information with continuous values random variables.
   
   Hint: maximize $I(X;Y)$ by showing that $h(Y|X)$ does not depend on $p(x)$, i.e. by maximizing $h(Y)$ as a function of $p(x)$.
   
   Additional hint: The capacity of this channel is also its zero error capacity.

3. Example 10.5 from Cover and Thomas: An exercise in computing the rate-distortion function. Follow the recipe of determining
   
   (a) the minimum distortion $D_{\text{min}}$ that must be incurred no matter how high a rate $R_{\text{max}}$ is permitted, and the corresponding rate, e.g. $R_{\text{max}} = H(X)$;
   
   (b) the maximum distortion $D_{\text{max}}$ beyond which $R = 0$ is possible;
   
   (c) the rate-distortion function

   $$ R(D) = \max_{p(\hat{x}|x): \sum_{x,\hat{x}} p(x)p(\hat{x}|x)d(x,\hat{x}) \leq D} I(X;\hat{X}) $$

   for $D_{\text{min}} \leq D \leq D_{\text{max}}$.

   For your own understanding, sketch $R(D)$ by evaluating it at a few intermediate points, e.g. $D_{\text{min}} + \delta$, $\frac{D_{\text{max}} - D_{\text{min}}}{2}$, $D_{\text{max}} - \delta$, or computing the $D$ that achieves $R(D) = \frac{R_{\text{max}}}{2}$.

4. Problem 10.2 from Cover and Thomas: Another exercise in computing the rate-distortion function.

   This time, when computing $R(D)$, note that since $d(0,1) = +\infty$, any finite distortion $D$ is achievable only with $p(0,1) = 0$.

5. Problem 10.4 from Cover and Thomas: Bonus for Ugrads, mandatory for Grads.

   This is a good illustration of how simplifying assumptions can be made “without loss of generality” when proving more general theorems.