

EE376A/Stats376A: Homework #1

Due on Thursday January 19, 5pm

You can hand in the homework either after class or deposit it, before 5 PM, in the EE376A drawer of the class file cabinet on the second floor of the Packard Building.

1. Random variables.

Let X be a (discrete) random variable with probability mass function (pmf) $p(0) = 1/2, p(1) = p(-1) = 1/6, p(2) = p(-2) = 1/12$.

- Plot the pmf of X . Compute $E[X]$.
- Let $Y = X^2$. Compute and plot the pmf q of Y .
- Compute $E[Y]$ in two ways: 1) by definition: $\sum_y q(y)y$, 2) shortcut: $\sum_x p(x)x^2$. Verify that they give the same answer. Is $E[Y]$ label-invariant?
- Let $Z = \log(1/p(X))$. Compute and plot the pmf r of Z .
- Compute $H(X) = E[Z]$ in two ways: 1) by definition: $\sum_z r(z)z$, 2) shortcut: $\sum_x p(x) \log(1/p(x))$. Verify that they give the same answer. Is $E[Z]$ label-invariant?
- Given a general random variable X with a pmf p and a random variable $Y = f(X)$, show that the shortcut always works, i.e.:

$$E[Y] = \sum_x p(x)f(x).$$

2. Multiple random variables.

Let X and Y be two binary random variables such that their joint pmf is $p(0,0) = p(1,1) = 1/3, p(0,1) = p(1,0) = 1/6$.

- Plot the joint pmf of X and Y .
- Let $Z = \log(1/p(X,Y))$. Compute and plot the pmf q of Z .
- Compute $H(X,Y) = E[Z]$ in two ways: 1) by definition: $\sum_z q(z)z$, 2) shortcut: $\sum_{x,y} p(x,y) \log(1/p(x,y))$. Verify that they give the same answer.
- Let $W = \log(1/p(Y|X))$. Compute and plot the pmf r of W .
- Compute $H(Y|X) = E[W]$ in two ways: 1) by definition: $\sum_w r(w)w$, 2) shortcut: $\sum_{x,y} p(x,y) \log(1/p(y|x))$. Verify that they give the same answer.
- Given two general random variables X and Y with a joint pmf p and a random variable $Z = f(X,Y)$, show that the shortcut always works, i.e.:

$$E[Z] = \sum_{x,y} p(x,y)f(x,y).$$

3. Example of joint entropy.

Let $p(x, y)$ be given by

	Y	
X	0	1
0	$\frac{1}{4}$	$\frac{1}{4}$
1	0	$\frac{1}{2}$

Find

- (a) $H(X), H(Y)$.
- (b) $H(X|Y), H(Y|X)$.
- (c) $H(X, Y)$.
- (d) $I(X; Y)$.

Draw a "pseudo" Venn diagram, labeling all above quantities.

4. Chain rule

Recall that the chain rule for entropy for two random variables is: $H(X, Y) = H(X) + H(Y|X)$.

- (a) Show that for three random variables X, Y, Z , a conditional chain rule holds:

$$H(Y, Z|X) = H(Y|X) + H(Z|X, Y).$$

- (b) Verify the chain rule for mutual information for three random variables starting from the chain rule for entropy. Recall that the chain rule for mutual information for three random variables X, Y, Z is:

$$I(X; Y, Z) = I(X; Y) + I(X; Z|Y).$$

5. Entropy of the sum of random variables

Suppose X and Y are two random variables taking values on $\{1, 2, \dots, n\}$. Let $Z = X + Y$ (note that Z takes value on $\{2, 3, \dots, 2n\}$).

- (a) Show that $H(Z|X) = H(Y|X)$.
- (b) Show that if X and Y are independent, then $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$.
Hint: Chain rule.
- (c) Give an example of X, Y such that $H(Z) < H(Y)$ and $H(Z) < H(X)$ (note that in this part X and Y are not necessarily independent).
- (d) Show that $H(Z) \leq H(X) + H(Y)$. Under what conditions the equality $H(Z) = H(X) + H(Y)$ holds?
Hint: You can start by proving $H(Z) \leq H(X, Y)$.

6. Independence vs. Conditional Independence

Suppose that X, Y, Z are three discrete random variables defined on $\{1, 2, \dots, n\}$ with a joint distribution $P_{X,Y,Z}$.

- (a) Show that X and Y are independent if and only if $I(X; Y) = 0$.
- (b) We define X and Y to be conditionally independent given Z , if

$$\forall x, y, z : P_{X,Y|Z}(x, y|z) = P_{X|Z}(x|z)P_{Y|Z}(y|z). \quad (1)$$

Show that X and Y are conditionally independent given Z if and only if $I(X; Y|Z) = 0$.

- (c) Does $I(X; Y|Z) = 0$ always imply $I(X; Y) = 0$? Either prove it or give a counterexample.
- (d) Does $I(X; Y) = 0$ always imply $I(X; Y|Z) = 0$? Either prove it or give a counterexample.

7. A Measure of Correlation

Define

$$\rho_{\text{MI}}(X, Y) = \frac{I(X; Y)}{\max\{H(X), H(Y)\}}.$$

In this problem, we prove some desired properties of ρ_{MI} as a correlation measure. Throughout this problem let X and Y be two discrete random variables defined on $\{1, 2, \dots, n\}$.

- (a) Show that ρ_{MI} is normalized, i.e. $0 \leq \rho_{\text{MI}}(X, Y) \leq 1$.
- (b) Show that $\rho_{\text{MI}}(X, Y) = 0$ if and only if X and Y are independent.
- (c) Show that $\rho_{\text{MI}}(X, Y) = 1$ if and only if there exists a bijection from X to Y .
- (d) Compare and contrast the above properties of $\rho_{\text{MI}}(X, Y)$ with the well-known Pearson correlation coefficient $\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$ where cov denotes the covariance and σ_X denotes the standard deviation of X .