Class 7 in-class problems, 18.05, Spring 2022

Concept questions

Concept question 1. Independence I

Roll two dice: X = value on first, Y = value on second

$X\backslash Y$	1	2	3	4	5	6	$p(x_i)$
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
$p(y_j)$	1/6	1/6	1/6	1/6	1/6	1/6	1

Are X and Y independent? 1. Yes 2. No

Concept question 2. Independence II

Roll two dice: X = value on first, T = sum

$X\backslash T$	2	3	4	5	6	7	8	9	10	11	12	$p(x_i)$
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0	1/6
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	1/6
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	1/6
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	1/6
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	1/6
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	1/6
$p(y_j)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36	1

Are X and Y independent? 1. Yes 2. No

Concept question 3. Independence III

Which of the following joint pdfs are the variables independent? (Each of the ranges is a rectangle chosen so that $\int\int f(x,y)\,dx\,dy = 1$.)

(i) $f(x,y) = 4x^2y^3$.

(ii) $f(x,y) = \frac{1}{2}(x^3y + xy^3)$.

(iii) $f(x,y) = 6e^{-3x-2y}$

(a) i **(b)** ii **(c)** iii **(d)** i, ii

(e) i, iii **(f)** ii, iii **(g)** i, ii, iii **(h)** None

Board questions

Problem 1. Joint distributions
Suppose X and Y are random variables and

- (X, Y) takes values in $[0, 1] \times [0, 1]$.
- the pdf is $f(x, y) = x + y$.

(a) Show $f(x, y)$ is a valid pdf.

(b) Visualize the event A = '$X > 0.3$ and $Y > 0.5$'. Find its probability.

(c) Find the cdf $F(x, y)$.

(d) Use the cdf $F(x, y)$ to find the marginal cdf $F_X(x)$ and $P(X < 0.5)$.

(e) Find the marginal pdf $f_X(x)$. Use this to find $P(X < 0.5)$.

(f) (New scenario) From the following table compute $F(3.5, 4)$.

$X \backslash Y$	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

Problem 2. Covariance and correlation
Flip a fair coin 11 times. (The tosses are all independent.)

Let X = number of heads in the first 6 flips

Let Y = number of heads on the last 6 flips.

Compute $\text{Cov}(X, Y)$ and $\text{Cor}(X, Y)$.

Problem 3. Even more tosses
Toss a fair coin $2n + 1$ times. Let X be the number of heads on the first $n + 1$ tosses and Y the number on the last $n + 1$ tosses.

Compute $\text{Cov}(X, Y)$ and $\text{Cor}(X, Y)$.

Extra

Discussion: Real-life correlations

- Over time, amount of ice cream consumption is correlated with number of pool drownings.

- In 1685 (and today) being a student is the most dangerous profession. That is, the average age of those who die is less than any other profession.
- In 90% of bar fights ending in a death the person who started the fight died.
- Hormone replacement therapy (HRT) is correlated with a lower rate of coronary heart disease (CHD).

Extra problem 1: Hospitals, binomial, CLT etc.
Here's one more problem. We won't do this in class.

- A certain town is served by two hospitals.
- Larger hospital: about 45 babies born each day.
- Smaller hospital about 15 babies born each day.
- For a period of 1 year, each hospital recorded the days on which more than 60% of the babies born were boys.

(a) Which hospital do you think recorded more such days?

(i) The larger hospital. (ii) The smaller hospital.

(iii) About the same (that is, within 5% of each other).

(b) Assume exactly 45 and 15 babies are born at the hospitals each day. Let L_i (resp., S_i) be the Bernoulli random variable which takes the value 1 if more than 60% of the babies born in the larger (resp., smaller) hospital on the i^{th} day were boys. Determine the distribution of L_i and of S_i.

(c) Let L (resp., S) be the number of days on which more than 60% of the babies born in the larger (resp., smaller) hospital were boys. What type of distribution do L and S have? Compute the expected value and variance in each case.

(d) Via the CLT, approximate the 0.84 quantile of L (resp., S). Would you like to revise your answer to part (a)?

(e) What is the correlation of L and S? What is the joint pmf of L and S? Visualize the region corresponding to the event $L > S$. Express $P(L > S)$ as a double sum.

Extra problem 2: Correlation
(a) Flip a coin 3 times. Use a joint pmf table to compute the covariance and correlation between the number of heads on the first 2 and the number of heads on the last 2 flips.

(b) Flip a coin 5 times. Use properties of covariance to compute the covariance and correlation between the number of heads on the first 3 and last 3 flips.