STAT1005 Week 9 Cheat Sheet

Joint Probability Distributions

- Is the probability of two different probability functions coming together to have a combined probability function for an outcome that combines the outcome of both functions.
  - For example you roll two dice, you want to compare the difference between number of dots showing on each dice, versus the number of dots showing on the dice with the highest number of dots
    - So you will have p(x) for the probability of any outcome having a difference of X
    - And p(y) for the probability of any outcome having Y dots showing on the dice with the most dots showing
    - The probability of a difference of x, and a total of y dots showing on the dice with the highest number of dots showing for a single roll will be given by combining the probabilities for each will be given by:
      - p(x, y)

Joint Probability Distributions of Discrete Random Variables

- ***Joint PMF***
  - p(x, y) = P(X = x and Y = y)
    - p(x, y) can be used as a joint pmf if:
      - p(x, y) $\geq$ 0 for all x and y
      - And $\sum_{x,y}$ p(x, y) = 1
  - Example of joint pmf for dice example
    - x: difference between dots showing
    - y: number of dots showing on dice with highest number of dots showing

x→ / y↓	0	1	2	3	4	5
1	1/36	0	0	0	0	0
2	1/36	2/36	0	0	0	0
3	1/36	2/36	2/36	0	0	0
4	1/36	2/36	2/36	2/36	0	0
5	1/36	2/36	2/36	2/36	2/36	0
6	1/36	2/36	2/36	2/36	2/36	2/36

  - So we can see that P(X = 0, Y = 1) = 1/36
    - You can put together all the variations to create a pmf, or:
  - ***Marginal Probability Mass Functions***
    - A pmf for one variable alone, created by summing p(x, y) over the values of the other variable
      - Marginal pmf of x: $p_x(x) = \sum_y p(x, y)$
      - Marginal pmf of y: $p_y(y) = \sum_x p(x, y)$
        - It is named marginal because the sums are just the values in a given row or column
          - Eg: In the dice example:
            - $P_x(X = 0)$ = 6(1/36) (or the sum of column x = 0)
            - P (Y = 6) = 1/36 + 5(2/36) (or the sum of row y = 6)
    - For A is any set of values consisting of pairs (x, y) values, such as {(x, y): x + y < 5}.
      - Then the probability that the random pair (X, Y) lies in A is found by summing the joint pmf over all pairs in A
      - $P[(X, Y) \in A] = \sum\sum_{(x,y)\in A} p(x, y) = \sum_x \sum_y p(x, y)$
        - Or $P[(X, Y) \in A] = p_x(x) + p_y(y)$

- ***Joint CDF***
  - $F(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x, v \leq y} p(u, v)$ for all x, y
  - Example joint cdf for dice example
    - x: difference between number of dots showing
    - y: number of dots showing on dice with highest number of dots showing
    - p (x): marginal pmf of x
    - p (y): marginal pmf of y

$y \downarrow$ \ $x \rightarrow$	**0**	**1**	**2**	**3**	**4**	**5**	**p (y)**
1	1/36	1/36	1/36	1/36	1/36	1/36	1/36
2	2/36	4/36	4/36	4/36	4/36	4/36	3/36
3	3/36	7/36	9/36	9/36	9/36	9/36	5/36
4	4/36	10/36	14/36	16/36	16/36	16/36	7/36
5	5/36	13/36	19/36	23/36	25/36	25/36	9/36
6	6/36	16/36	24/36	30/36	34/36	36/36	11/36
p (x)	6/36	10/36	8/36	6/36	4/36	2/36	-

- ***Independence of Random Variables***
  - If information about one event can give us new information about another event, then the two events are not independent
    - In the dice example we can see that:
      - $P(X = 1|Y = 2) = 2/36$ and $P(X = 0| Y = 2) = 1/36$
        - That $P(X = 1)$ is twice as likely as $P(X = 0)$ for the same value of Y shows that there is dependence between the two variables
  - As we saw previously one way of defining the independence of two events is via the condition of independence:
    - $P(A \cap B) = P(A).P(B)$
  - For two random variables X and Y, they are said to be independent if for any values of x and y,
    - $p(x, y) = p_x(x).p_y(y)$
      - If this is not satisfied for all (x, y), then X and Y are said to be dependent
        - The same is true for continuous random variables
- ***Expected Values***
  - We have seen previously that any function h(x) of a single random variable X is itself a random variable
    - $\mu_{h(x,y)} = E[h(x)] = \sum h(x).p(x)$
  - So, for X and Y jointly distributed discrete random variables with pmf (x, y). The expected values is given using the expected value function h(x, y):
    - $\mu_{h(x,y)} = E[h(x, y)] = \sum_x \sum_y h(x, y).p(x, y)$
- ***Covariance***
  - When two random variables are not independent, we will want to see how strongly they are related to each other, or how much they "co-vary"
  - The covariance between two discrete random variables is:
    - $Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$
      - $= \sum_x \sum_y (x - \mu_X)(y - \mu_Y).p(x, y)$
      - $= E[X, Y] - E[X]E[Y]$
    - If X and Y move together positively or negatively (same direction) then the covariance will be positive
    - If X and Y move away from each other (different directions) then the covariance will be negative
    - If X and Y are not strongly related, they will cancel each other out and the covariance will near 0

- **Correlation**
  - Covariance tells us about the direction of the relationship between X and Y, but not the strength of the relationship
  - To scale this relationship we use the correlation coefficient:
    - $\rho(x,y) = Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$
  - *Additional considerations*:
    - 1. If X and Y are independent, then Corr(X, Y) = 0, but Corr(X, Y) does not imply independence
    - 2. Corr(X, Y) = 1 or -1 if and only if Y = aX + b for some numbers a and b with a ≠ 0
      - a) Correlation is a measure of the strength of the linear relationship between X and Y
      - b) Two variables can be uncorrelated yet highly dependent because of a strong nonlinear relationship

Joint Probability Distributions of Continuous Random Variables

- Essentially the same as dealing with joint probability distributions of discrete random variables, but:
  - Discrete sets and replaced by continuous intervals
  - Joint pmf is replaced by a joint pdf
  - Sums are replaced by integrals
- ***Joint pdf***
  - For the probability the observed value of a continuous random variable X lies in set A is obtained by integrating the pdf f(x) over the set A, where A is the interval [a, b]:
    - $P(a \leq X \leq b) = \int_a^b f(x)\, dx$
      - You do the same thing to find the probability of a the pair (X, Y) falls in set A where x falls with a and b and y falls within c and d, where A is defined as the region {(x, y): $a \leq x \leq b$, $c \leq y \leq d$}, which can be denoted as [a, b] x [c, d]:
  - So for continuous random variables X and Y f(x, y) is the joint probability density function if for any two dimensional set A:
    - $P[(X, Y) \in A] = \iint_A f(x, y)\, dx\, dy$
      - $P[(a < x < b, c < y < d)] = \int_a^b \int_c^d f(x, y)\, dx\, dy$
  - f(x, y) is a joint pdf if:
    - 1. $f(x, y) \geq 0$
    - 2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$
  - These double integrals is essentially double adding, called iterated integrals
    - 1. Inner integral: hold x constant, integrate over y
    - 2. Outer integral: hold y constant, integrate over x
    - Which step is integrating for x or y doesn't matter
- ***Marginal pdfs***
  - The marginal pdfs for X and Y
    - $f(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$, for $-\infty < x < \infty$
    - $f(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$, for $-\infty < y < \infty$
- ***Independence of Continuous Random Variables***
  - Two continuous random variables are said to be independent if for every pair of x and y values:
    - $f(x, y) = f_x(x).f_y(y)$
    - AND the region of positive density has side parallel to the axis
- ***Expected Value***
  - The expected value for a continuous pdf f(x, y) with the expected value function h(x, y):
    - $\mu_{h(x,y)} = E[h(x, y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x, y).f(x, y)\, dx\, dy$
  - *Properties of Expected Value:*
    - Linearity of expectation:
      - For any functions h_1, h_2, and any constants a_1, a_2, b
        - $E[(a_1h_1(x, y) + a_2h_2(x, y) + b] = a_1E[h_1(x, y)] + a_2E[h_2(x, y)] + b$
    - Special form of h(x, y)
      - If X and Y are independent and $h(X, Y) = f_x(X).f_y(Y)$ then:
        - $E[h(X, Y)] = E[f_x(X).f_y(Y)] = E[f_x(X)].E[f_y(Y)]$

- ***Covariance***
  - The covariance between two continuous random variables X and Y:
    - $Cov(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$
      - $= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y).f(x, y)\, dx\, dy$
      - $= E[X, Y] - E[X]E[Y]$
  - *Properties of Covariance*
    - 1. $Cov(X, Y) = Cov(Y, X)$
    - 2. $Cov(X, X) = Var(X)$
    - 3. Shortcut formula:
      - $Cov(X, Y) = E(XY) - \mu_x \mu_y$
    - 4. Distributive property: For any random variable Z and any constants a, b, c
      - $Cov(aX + bY + c, Z) = aCov(X, Z) + bCov(Y, Z)$
- ***Correlation***
  - The correlation coefficient of X and Y, denoted by Corr(X, Y) or ρ_{xy} is defined as:
    - $Corr(X, Y) = \rho_{xy} = \dfrac{Cov(x, y)}{\sigma_x \sigma_y}$