The Multinomial Distribution

1. A bag of 20 candies contains candy pieces of an assortment of colors. Suppose that the color of each candy piece is determined by the following probabilities, independent of the colors of other candy pieces.

Color	Red	Orange	Yellow	Green	Blue	Indigo	Violet
Probability	.2	.1	.15	.2	.1	.1	.15

   Let R, O, Y, G, B, I, V denote the number of candy pieces in a particular bag of the respective colors.

   (a) Explain why the vector (R, O, Y, G, B, I, V) has multinomial distribution, and explicitly write down the parameters of this distribution.

   (b) What is the probability that a single bag of candy contains at least one candy of each color?

   (c) My son Oliver hasn't learned the color Indigo, and so treats Indigo and Violet candies as indistinguishable. Let P denote the number of purple (i.e. Indigo or Violet) candies. What is the marginal distribution of P? What is the joint distribution of (R, O, Y, G, B, P)? Be sure to specify the parameters for each distribution.

   (d) Suppose, due to a complicated candy machine malfunction, that candy bags are produced with no Red candies. What is the conditional distribution of (O, Y, G, B, I, V) given this malfunction?

2. Suppose a bowl contains 700 pieces of candy, with 100 pieces of each color (Red, Orange, Yellow, Green, Blue, Indigo, Violet). Moreover, suppose you scoop out 20 pieces of candy at random, so that every subset of size 20 is equally likely. Let R, O, Y, G, B, I, V denote the number of candies in a the scoop of the respective colors.

   (a) Explain why the distribution of (R, O, Y, G, B, I, V) is **not** multinomial.

   (b) Find a formula for joint PMF of (R, O, Y, G, B, I, V). *Hint: It may be helpful to first think of a slightly different scenario where there are only 2 colors of candy.*

   (c) Explain how to modify the sampling procedure so the resulting joint distribution **IS** Multinomial.