DSE 210 (Probability and Statistics Using Python)
Final Exam, due Monday 3/15 5:00pm

---

1. a) Pr(first care is a heart) = 1/4
   b) Pr(first card is not a spade) = 1 − (1/4) = 3/4
   c) Pr(first and second card are the same colour) = 25/51
   d) Pr(fifth card is red) = 1/2
   e) Pr(first card is ace | first card is a heart) = 1/13
   f) Pr(second card is an ace | first card is an ace) = 3/51

2. Only the second option is independent.

3. a) Each permutation is equally likely, so the probability of this event is $\frac{1}{5!} = 1/120$
   b) For some fixed value, the probability all die show this value is $(1/6)^3 = 1/216$. Taking the union over all 6 outcomes, the probability of this event is: $6/216 = 1/36 = 0.028$
   c) For a binary outcome that occurs with probability p, the expected wait time for the first success is $1/p$. Therefore, we would expect to run into our worst enemy on the $1/0.2 = 5$ trip to the gym. So the expected number of trips *before* running into them is 4 (but 5 is also an acceptable answer).
   d) We can compute this using Bayes rule:

$$\Pr(\text{female} \mid \text{lefty}) = \frac{\Pr(\text{lefty} \mid \text{female})\Pr(\text{female})}{\Pr(\text{lefty})} = \frac{0.1 \cdot 0.6}{0.6 \cdot 0.1 + 0.4 \cdot 0.2} = 6/14 = 0.428$$

4. a) Using the law of total probability:

$$\Pr(\text{rash}) = \Pr(\text{rash} \mid \text{bite})\Pr(\text{bite}) + \Pr(\text{rash} \mid \text{no bite})\Pr(\text{no bite}) = 1 \cdot \frac{1}{9} + \frac{1}{8} \cdot \frac{8}{9} = \frac{2}{9} \approx 0.22$$

   b) Using Bayes rule:

$$\Pr(\text{bite} \mid \text{rash}) = \frac{\Pr(\text{rash} \mid \text{bite})\Pr(\text{bite})}{\Pr(\text{rash})} = \frac{1/9}{2/9} = \frac{1}{2}.$$

5. a) Using the definition of expectation:

$$\mathbb{E}[X] = \sum_{i=1}^{6} \Pr(X = i) \cdot i = \frac{1}{12} + \frac{2}{12} + \frac{3}{12} + \frac{4}{4} + \frac{5}{4} + \frac{6}{4} = 4.25$$

   b) Using the definition of variance:

$$\texttt{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \sum_{i=1}^{6} \Pr(X = i) \cdot i^2 - 4.25^2 = 20.4167 - 18.0625 = 2.35$$

   c) Let X_i be a random variable denoting the outcome of the i-th roll. By linearity of expectation and independence:

$$\mathbb{E}[Z] = \mathbb{E}\left[\frac{1}{100}\sum_{i=1}^{100} X_i\right] = \frac{1}{100}\sum_{i} \mathbb{E}[X_i] = \mathbb{E}[X_i] = \mathbb{E}[X] = 4.25$$

d) Using the fact that the rolls are independent and identically distributed:

$$\texttt{Var}(Z) = \texttt{Var}\left(\frac{1}{100}\sum_{i=1}^{100} X_i\right) = \frac{100}{100^2}\texttt{Var}(X_i) = \frac{1}{100}\texttt{Var}(X) = 0.0235$$

6. a) By linearity of expectation:

$$\mathbb{E}[W] = \mathbb{E}[X - Y + Z] = \mathbb{E}[X] - \mathbb{E}[Y] + \mathbb{E}[Z] = 1 - 0 + 2 = 3$$

b) By independence:

$$\texttt{Var}(W) = \texttt{Var}(X - Y + Z) = \texttt{Var}(X) + \texttt{Var}(Y) + \texttt{Var}(Z) = 14 + 4 + 9 = 29.$$

Note that we do not subtract variances in a linear combination. We can think of $\texttt{Var}(X - Y)$ as $\texttt{Var}(X) + \texttt{Var}(-1 \cdot Y) = \texttt{Var}(X) + \texttt{Var}(Y)$ to see this. More intuitively, adding together random variables should not decrease their randomness.

7. The mean vector is simply the mean of the respective variables:

$$\mu = \begin{pmatrix}\mathbb{E}[X_1] \\ \mathbb{E}[X_2]\end{pmatrix} = \begin{pmatrix}0 \\ 0.5\end{pmatrix}$$

The covariance matrix is given by:

$$\Sigma = \begin{pmatrix}\texttt{Var}(X_1) & \texttt{Cov}(X_1, X_2) \\ \texttt{Cov}(X_1, X_2) & \texttt{Var}(X_2)\end{pmatrix}$$

First let us compute the variances:

$$\texttt{Var}(X_1) = \mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 = 1 - 0 = 1.$$

Similarly:

$$\texttt{Var}(X_2) = \mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2 = 1 - 0.25 = 0.75.$$

Now recall:

$$\texttt{Cov}(X_1, X_2) = \texttt{Std}(X_1)\texttt{Std}(X_2)\texttt{Corr}(X_1, X_2) = 0.2165.$$

So:

$$\Sigma = \begin{pmatrix}1 & 0.2156 \\ 0.2156 & 0.75\end{pmatrix}$$

8. a) The MLE for the rate parameter of the Poisson is just the sample mean rate:

$$\hat{\lambda} = \frac{5 + 3 + 1 + 0 + 0 + 1 + 2 + 4 + 3 + 4}{10} = \frac{23}{10} = 2.3$$

b) Since $\texttt{Cov}(X_1, X_2) = 0$ this Gaussian distribution is aligned with the canonical axes in $\mathbb{R}^2$ and is symmetric. Therefore exactly half the mass of the distribution lies above $\mathbb{E}[X_2]$, and so $\Pr(X_2 \geq 4) = 1/2$.

c) Let p_v be the multinomial parameter associated with $v \in V$. The MLE for p is simply the empirical probability of observing v. With Laplace smoothing:

$$\hat{p}_v = \frac{N_v + 1}{N + |V|},$$

where N_v is the number of times v appears in the document and $N = \sum_{v \in V} N_v$. Thus:

v	beside	dog	cat	another
p_v	2/11	4/11	2/11	3/11

9. A 95% CI is given by: $[\mu - 1.96\sigma, \mu + 1.96\sigma]$. We can estimate μ by the sample mean: $\hat{\mu} = 12.2$ and $\hat{\sigma} = \sigma/\sqrt{N} = 5.4/10 = 0.54$. Therefore, the 95% CI is: $[11.14, 13.26]$. It would also be acceptable to use 2σ in which case the CI is: $[11.12, 13.28]$. Note though that it's better to refer to the latter case as a 2σ confidence interval. That way there's no ambiguity.

10. a) Let μ_1 be the mean SAT score from Genius Academy and let μ_2 be the mean score from the local high school. The null hypothesis is: $h_0 : \mu_1 = \mu_2$ which we will test against the alternative $h_a : \mu_1 \neq \mu_2$. We could also do a one tailed test against $\mu_1 > \mu_2$, since the name "Genius Academy" seems to suggest they think their scores should be higher.

    b) The z-statistic for a difference in means is given by:

$$z = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

    Plugging in the numbers for this problem, we find:

$$\sigma_1 = \frac{150}{10} = 15\, \sigma_2 = \frac{200}{10} = 20 \Rightarrow z = \frac{1930 - 1860}{25} = 2.8$$

    c) The p-value associated with this z-score is:

$$p = 2(1 - \Phi(2.8)) = 0.0051.$$

    This is strong evidence against the null so we can conclude that the observed difference is significant. If we want to do a 1-tailed test the p-value is:

$$p = 1 - \Phi(2.8) = 0.0026$$

11. a) A 95% confidence interval is given by $p \pm 1.96\sigma$. We are given an estimate of $\hat{p}$ which leads to an estimate for the standard error of:

$$\hat{\sigma} = \sqrt{\frac{\hat{p}(1-\hat{p})}{N}} = 0.049.$$

    Therefore, the confidence interval is: $[0.304, 0.496]$. It would also be acceptable to use 2 standard errors which would yield a confidence interval of $[0.302, 0.498]$.

    b) We want the width of the confidence interval to be at most $\pm\epsilon$. Therefore, we want:

$$1.96\sigma \leq \epsilon$$
$$1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{N}} \leq \epsilon$$
$$N \geq \hat{p}(1-\hat{p}) \left(\frac{1.96}{\epsilon}\right)^2$$

    Plugging in the values from this problem we find:

$$N \geq 9219.8$$

    So we should use a sample size of 9220. Approximating the 95% CI with 2 standard errors yields: $N = 9600$.