Erreta

Q1

(c) ... size of the randomly chosen group. Let E[Y] $= \mu$ and ...

Partial solution to HW 6

Q1

Checkout out chapter 5 Q5 solution in Janko Graver's book.

Q2

(Panchenko Exercise 1.5.2) There could be many, my solution use shifted normalized Poisson Set discrete random variable X. We define the pmf f_X to be

$$f_X(x) \begin{cases} 1 - e^{-10}, & \text{if } x = 0; \\ e^{-10} e^{\lambda} \frac{\lambda^{(x-1)}}{(x-1)!}, & \text{if } x \in \mathbb{N}. \end{cases}$$

Then, $P(X > 0) = e^{-10}$ is easy to check. Now we have to set the correct λ. By some calculation you can see that

$$\begin{aligned} E[X] &= \sum_{x=1}^{\infty} ((x-1)+1) e^{-10} e^{\lambda} \frac{\lambda^{(x-1)}}{(x-1)!} \\ &= e^{-10} \underbrace{\sum_{x=0}^{\infty} x e^{\lambda} \frac{\lambda^{x}}{(x)!}}_{\text{expectation of Poisson}(\lambda)} + e^{-10} \\ &= e^{-10}(\lambda + 1) \end{aligned}$$

Hence, set $\lambda = e^{20} - 1$ we will have $E[X] = e^{10}$.

$\square$

Q3

This is the same as the exercise in RC4 exercise for the indicator method. The number of pairs of animals alive, N, is not an easy random variable to compute. Therefore, we may consider "when will N increase?". This is easier, N increase 1 if a pair of animals is alive. Hence, the indicator we consider is $1_{A_i}, i = 1, \ldots, n$, where A_i means the event that the ith pair of animal is alive.

Then, we have $E[1_{A_i}] = P(A_i) = \binom{2(n-1)}{m} / \binom{2n}{m}$

Consequently, by linearity of expectation,
$E[N] = \sum_{i=1}^{n} E[1_{A_i}] = nP(A_1) = \frac{(2n-m)(2n-m-1)}{2(2n-1)}$ pairs.

$\square$

Q4, Q5

these two will need some inclusion-exclusion.

Q6

(a) Calculate $E[X^2] = p$, and use $\text{Var}(X) = E[X^2] - E[X]^2$.

(b) We did that in recitation class. If you forget, try to use the method introduced in handout Chapter 3 page 14 of from prof. Sheu.

(c) Recall that $E[aX + b] = aE[X] + b$ due to the linearity of expectation. We have $Var(aX + b) = E[(aX + b - (aE[X] + b))^2] = E[a^2(X - E[X])^2] = a^2Var(X)$

Q7

First, we can eyeball $E[Y] = 0$ because $(X_i)_{i=1}^n$ are i.i.d and we have $E[X_i] = 0$. Hence, what is left is to calculate $E[Y^2]$. Before we dive in to calculation, again we have to observe which terms will not be zero? Those with first order X_i will be zero, e.g. $E[X_1^2X_2X_3], E[X_1X_2X_3X_4]$ are zero. Hence, we only have to consider terms like $E[X_1^2X_2^2]$ that have only second order moment. Also $E[X_i^2X_j^2] = 1,$ for any $1 \leq i < j \leq n$.

$$E[Y^2] = \sum_{k<l}\sum_{i<j} E[X_iX_jX_kX_l] = \sum_{i<j} E[X_i^2X_j^2] = \frac{n(n-1)}{2} \cdot E[X_1^2X_2^2] = \frac{n(n-1)}{2}$$

$\square$

Q8

Check out Chapter 3 handout page 14 by prof. Sheu.

Q9

This is a problem using the fact that $E[XY] = E[X]E[Y]$ if X, Y are independent.

- Show that $E[XY] = E[X]E[Y]$ if X, Y are independent.

  This is true for both discrete and continuous random variable, but due to our coverage, we show the discrete version of it.

$$\begin{aligned} E[XY] &= \sum_{x,y} xyP(X = x, Y = y) \underbrace{=}_{\text{independent}} \sum_{x,y} xyP(X = x)P(Y = y) \\ &= \sum_{x} xP(X = x) \sum_{y} yP(Y = y) = E[X]E[Y] \end{aligned}$$

Hence we have $E[XY] = E[X]E[Y]$. To further simplify our computation, we should note that X_1^2, X_2^2 are also independent (checkout recitation 3 review exercise, and take $g(x), h(x)$ both be x^2) With these two results we may proceed to calculate $\text{Var}(X_1X_2)$

$$\text{Var}(X_1X_2) = E[(X_1X_2)^2] - E[X_1X_2] = E[X_1^2]E[X_2^2] - E[X_1]E[X_2]$$

which is $(\sigma_1^2 + \mu_1^2)(\sigma_2^2 + \mu_2^2) - \mu_1^2\mu_2^2$.

If you want to go through the calculation, here it is.

$$\begin{aligned}
\text{Var}(X_1X_2) &= E[(X_1X_2 - E[X_1X_2])^2] = E[(X_1X_2 - E[X_1]E[X_2])^2] \\
&= \sum_{x_1,x_2} (x_1x_2 - E[X_1]E[X_2])^2 P(X_1 = x_1, X_2 = x_2) \\
&= \sum_{x_1,x_2} (x_1x_2 - E[X_1]E[X_2])^2 P(X_1 = x_1)P(X_2 = x_2) \\
&\qquad \text{(independent)} \\
&= \sum_{x_2} P(X_2 = x_2) \sum_{x_1} (x_1^2x_2^2 - 2x_1x_2E[X_1]E[X_2] + E[X_1]^2E[X_2]^2)P(X_1 = x_1) \\
&\qquad \text{(Expand the terms and separate the summation)} \\
&= \sum_{x_2} P(X_2 = x_2)(E[X_1^2]x_2^2 - 2E[X_1]^2x_2E[X_2] + E[X_1]^2E[X_2]^2) \\
&= E[X_1^2]E[X_2^2] - 2E[X_1]^2E[X_2]^2 + E[X_1]^2E[X_2]^2 \\
&= E[X_1^2]E[X_2^2] - E[X_1]^2E[X_2]^2 = (\sigma_1^2 + \mu_1^2)(\sigma_2^2 + \mu_2^2) - \mu_1^2\mu_2^2
\end{aligned}$$

The same as we had earlier.

Q10

Q11

Use the method of indicators, for $i \neq j$, we can write

$$\begin{aligned}
\mathbb{E}(X_{e(i)}X_{e(j)}) &= \mathbb{E}(X_{e(i)}X_{e(j)}\mathbb{1}_{\{e(i)\neq e(j)\}} + X_{e(i)}X_{e(j)}\mathbb{1}_{\{e(i)=e(j)\}}) \\
&= \mathbb{E}(X_{e(i)}X_{e(j)}\mathbb{1}_{\{e(i)\neq e(j)\}}) + \mathbb{E}(X_{e(i)}X_{e(j)}\mathbb{1}_{\{e(i)=e(j)\}}).
\end{aligned}$$

Adapt the indicator method again, you can calculate

$$\begin{aligned}
\mathbb{E}(X_{e(i)}X_{e(j)}\mathbb{1}_{\{e(i)\neq e(j)\}}) &= \sum_{k\neq i;\, l\neq j;\, k\neq l} \mathbb{E}(X_{e(i)}X_{e(j)}\mathbb{1}_{\{e(i)=k,\, e(j)=l\}}) \\
&= \sum_{k\neq i;\, l\neq j;\, k\neq l} \mathbb{E}(X_kX_l\mathbb{1}_{\{e(i)=k,\, e(j)=l\}}) \\
&= \sum_{k\neq i;\, l\neq j;\, k\neq l} \mathbb{E}(X_k)\mathbb{E}(X_l)\mathbb{E}(\mathbb{1}_{\{e(i)=k\}})\mathbb{E}(\mathbb{1}_{\{e(j)=l\}}) \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(X_{e(i)}X_{e(j)}\mathbb{1}_{\{e(i)=e(j)\}}) &= \sum_{k\neq i;\, k\neq j} \mathbb{E}(X_{e(i)}X_{e(j)}\mathbb{1}_{\{e(i)=k,\, e(j)=k\}}) \\
&= \sum_{k\neq i;\, k\neq j} \mathbb{E}(X_k^2\mathbb{1}_{\{e(i)=k,\, e(j)=k\}}) \\
&= \sum_{k\neq i;\, k\neq j} \mathbb{E}(X_k^2)\mathbb{E}(\mathbb{1}_{\{e(i)=k\}})\mathbb{E}(\mathbb{1}_{\{e(j)=k\}}) \\
&= \sum_{k\neq i;\, k\neq j} 1\cdot\mathbb{P}(e(i)=k)\mathbb{P}(e(j)=k) \\
&= \sum_{k\neq i;\, k\neq j} \frac{1}{(n-1)^2} = \frac{n-2}{(n-1)^2}.
\end{aligned}$$

Hence $\mathbb{E}(X_{e(i)}X_{e(j)}) = \frac{n-2}{(n-1)^2}$ for $i\neq j$.

When $i=j$, $\mathbb{E}\big(X_{e(i)}^2\big) = \sum_{k\neq i}\mathbb{E}(X_{e(i)}^2\mathbb{1}_{\{e(i)=k\}}) = \sum_{k\neq i}\mathbb{E}(X_k^2\mathbb{1}_{\{e(i)=k\}})$. Use the independence calculation again, you can see that $\mathbb{E}\big(X_{e(i)}^2\big) = (n-1)\cdot 1\cdot\frac{1}{(n-1)} = 1$.

Then the variance can be computed as

$$\begin{aligned}
Var(X_{e(1)}+\cdots+X_{e(n)}) &= \sum_{i,j} Cov(X_{e(i)}, X_{e(j)}) = \sum_{i,j}\mathbb{E}(X_{e(i)}X_{e(j)}) - \mathbb{E}(X_{e(i)})\mathbb{E}(X_{e(j)}) \\
&= \sum_{i,j}\mathbb{E}(X_{e(i)}X_{e(j)}) = \sum_{i=j} 1 + \sum_{i\neq j}\frac{n-2}{(n-1)^2} = n + \frac{n(n-2)}{n-1}.
\end{aligned}$$

You can check that $\mathbb{E}(X_{e(i)}) = 0$ with the similar method. $\square$