Week 4

Exercise 41 In order to get more money, an airline uses the fact that only 90% of people who book a flight, show up at the airport. On a particular flight with 300 seats, the airline decided to book 324 people.

(1) Assuming independence, what is the chance that the flight will be overbooked? (More than 300 people will show up.)

(2) Assuming people travel in pairs, what is the chance that the flight will be overbooked? (Assume the probability of showing up doesn't change)

(Use normal approximations and then redo with skew normal approximations.)
Answer. (1) We know that $n = 324$, $p = 0.9$, $\mu = 291.6$, and $\sigma = \sqrt{291.6 \cdot 0.1} = 5.4$. We have:

$$1 - \Phi\left(\frac{301 - \frac{1}{2} - 291.6}{5.4}\right) = 1 - \Phi(1.65) = 0.0495$$

For skew normal we let $z = \frac{300+\frac{1}{2}-291.6}{5.4} = 1.65$

$$\begin{aligned}1 - \left(\Phi(1.65) - \frac{1}{6}\operatorname{skw}(324, 0.9)(z^2-1)\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}z^2}\right) &= 1 - \left(0.9505 - \frac{1}{6\sqrt{2\pi}}\frac{1-2\cdot 0.9}{\sigma}(1.7225)e^{-1.36125}\right) \\ &= 1 - (0.9505 - 0.0665 \cdot (-0.14815) \cdot 1.7225 \cdot 0.25634) \\ &= 1 - (0.9505 + 0.0044) \\ &= 0.0451\end{aligned}$$

(2) We suppose all seats come in pairs. So we have $n = 162$, $p = 0.9$, $\mu = 145.8$, $\sigma = 3.82$ giving us:

$$1 - \Phi\left(\frac{151 - \frac{1}{2} - 145.8}{3.82}\right) = 1 - \Phi(1.23) = 0.1093$$

For skew normal we let $z = \frac{150+\frac{1}{2}-145.8}{3.82} = 1.23$.

$$\begin{aligned}1 - \left(\Phi(1.23) - \frac{1}{6}\operatorname{skw}(162, 0.9)(z^2-1)\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}z^2}\right) &= 1 - \left(0.8907 - \frac{1}{6\sqrt{2\pi}}\frac{1-2\cdot 0.9}{\sigma}(0.5129)e^{-0.75645}\right) \\ &= 1 - (0.8907 - 0.0665 \cdot (-0.2094) \cdot 0.5129 \cdot 0.46933) \\ &= 1 - (0.8907 + 0.0034) \\ &= 0.1127\end{aligned}$$

□

Exercise 42 Suppose that we roll a fair die 100 times. We know that the probability of getting 25 or more rolls of six happens with probability 0.022. If you rolled the fair die 100 times every day for a year (365 days), find the chance that you have 25 or more sixes:

(1) At least once

(2) At least twice

(Use Poisson approximation)
Answer. Note that $\mu = np = 365 \cdot 0.022 = 8.03$ Recall that the Poisson approximation is given by $P(k) = e^{-\mu}\frac{\mu^k}{k!}$.

(1)

$$P(\geq 1) = 1 - P(0) \approx 1 - e^{-\mu}\frac{\mu^0}{0!} = 1 - e^{-\mu} = 0.999674$$

(2)
$$P(\geq 2) = 1 - P(0) - P(1) \approx 0.99674 - e^{-\mu}\frac{\mu^1}{1!} = 0.997060$$

□

Exercise 43 A hat contains 1000 random numbers in it. Only two of the numbers are prime and the rest are composite (not prime).

(1) What is more likely to occur if we pull 1000 numbers from the box with replacement:
- Less than 2 prime numbers are pulled
- Exactly 2 prime numbers are pulled
- More than 2 prime numbers are pulled

(2) If we repeat the 1000 pulls just like the previous question, but this time do it twice in a row (so, do 1000 pulls with replacement; and then do a second 1000 pulls with replacement); what are the chances that we get the same number of prime numbers in both 1000 pulls?

(Use Poisson approximation)
Answer. We know that $\mu = np = 1000 \cdot \frac{2}{1000} = 2$.

(1) We use Poisson approximation to make things easier.

$$P(< 2) = P(0) + P(1) \approx e^{-\mu}\frac{\mu^0}{0!} + e^{-\mu}\frac{\mu^1}{1!} = e^{-2}(1+2) = 0.406006$$
$$P(2) \approx e^{-\mu}\frac{\mu^2}{2!} = 0.270671$$
$$P(> 2) = 1 - P(< 2) - P(2) \approx 0.323323$$

So we have the best chance of getting strictly less than 2 prime numbers.

(2) Here we ask when two series of pulls have the same number of prime numbers.

$$\begin{aligned}
P(\text{Both have same \# of prime \#}) &= \sum_{k=0}^{\infty} P(\text{Both have } k \text{ prime numbers}) \\
&\approx \sum_{k=0}^{\infty} e^{-\mu}\frac{\mu^k}{k!} \cdot e^{-\mu}\frac{\mu^k}{k!} \quad \text{(b/c they are independent)} \\
&= \sum_{k=0}^{\infty} e^{-2}e^{-2}\frac{2^k \cdot 2^k}{k!k!} \quad \text{(plug in } \mu = 2\text{)} \\
&= e^{-4}\sum_{k=0}^{\infty}\frac{4^k}{k!k!} \\
&= 0.018315639\,(1 + 4 + 4 + 1.77777 + 0.44444 + 0.07111 + 0.00790 + 0.00064 + \ldots) \\
&= 0.2070
\end{aligned}$$

Since we're approximating, we only need a few terms and don't care about the rest.

□

Exercise 44 Suppose you have a normal deck of 52 cards and you deal out 3 cards randomly. Recall that there are 26 red cards and 26 black cards.

(1) What's the probability that the first card is red and the second two are black?

(2) What's the probability that exactly one card is red?

(3) What's the probability that at least one card is red?

Answer. Our large population here is $N = 52$ and we know that $R = 26$ are red and $B = 26$ are black. So if we pull out 3 then $n = 3$. By our formulas for sampling without replacement we have:

$$P(r \text{ red in } n \text{ pulls}) = \binom{n}{r} \frac{(R)_r (B)_{n-r}}{(N)_n}$$

(1) We know that the first one is red and the last two are black. In this case, we are looking at just one combination instead of all. So we have:

$$\frac{26}{52} \cdot \frac{26}{51} \cdot \frac{25}{50} = \frac{13}{102} = 0.12745$$

(2) If we want to know the probability that exactly 1 are red, we can look at all combinations of the previous example. The red can be in the first, second or third place so we have:

$$3 \cdot \frac{13}{102} = \frac{13}{34} = 0.3823529$$

Another way to do this is to use our formula from above:

$$\binom{3}{1} \frac{(26)_1 \cdot (26)_2}{(52)_3} = \frac{3!}{1!2!} \cdot \frac{\frac{26!}{25!} \cdot \frac{26!}{24!}}{\frac{52!}{49!}} = \frac{3}{1} \cdot \frac{26 \cdot 26 \cdot 25}{52 \cdot 51 \cdot 50}$$

Which is the same equation as above.

(3) We want to look at the chances that none of the cards are red and remove that option. So we have:

$$1 - \frac{26}{52} \cdot \frac{25}{51} \cdot \frac{24}{50} = 1 - \frac{2}{17} = \frac{15}{17} = 0.8823529$$

□

Exercise 45 In the magical forest, out of every $100,000$ creatures, 40% are unicorns and 60% are horses. Say we pull a random sample of 100 creatures without replacement. What's the exact probability that there are at least 45 unicorns in the sample? (Don't solve. Keep it as a summation) Using normal approximation, give an approximation of the probability instead.

Answer. The exact chance is given by

$$\begin{aligned} P(45) + P(46) + \cdots + P(100) &\approx \binom{100}{45} \frac{(40,000)_{45}(60,000)_{55}}{(100,000)_{100}} + \cdots + \binom{100}{100} \frac{(40,000)_{100} \cdot (60,000)_0}{(100,000)_{100}} \\ &= \sum_{k=45}^{100} \binom{100}{k} \frac{(40,000)_k \cdot (60,000)_{n-k}}{(100,000)_{100}} \end{aligned}$$

The approximation is given by: $\mu = np = 100 \cdot 0.4 = 40$ and $\sigma = \sqrt{\mu(1-p)} = \sqrt{40 \cdot 0.6} = 4.9$ Then:

$$1 - \Phi\left(\frac{45 - \frac{1}{2} - 40}{4.9}\right) = 1 - \Phi(0.9184) = 0.1788$$

□

Exercise 46 Suppose you go to a small fair with 9 of your friends and you each purchase 10 tickets in the fair's raffle. Later on, you found out, your group of 10 people were the only people to purchase tickets for the raffle! Out of the 100 tickets bought, 3 of them will be chosen randomly as winning tickets.

(1) What are the chances one person will get all 3 winning tickets?

(2) What are the chances that 3 different people will win something?

(3) What are the chances that 2 different people will win?
Answer. This problem is a little complicated so let's break it completely down. The sample space consists of *sets* of three numbers. Since we're choosing three numbers out of 100, we know that there are $\binom{100}{3}$ total number of subsets.

(1) If one person wins all three, then the subset of three tickets that was chosen is also a subset of three tickets from the persons ten tickets. So it is one of $\binom{10}{3}$ possibilities. So the probability of any one particular person winning is:

$$\frac{\binom{10}{3}}{\binom{100}{3}} = \frac{\frac{10\cdot 9\cdot 8}{3\cdot 2}}{\frac{100\cdot 99\cdot 98}{3\cdot 2}} = \frac{10\cdot 3\cdot 4}{50\cdot 33\cdot 98} = \frac{2}{5\cdot 11\cdot 49}$$

But, there are ten people and so there are ten different ways of this occurring. Therefore we have a total probability of:

$$10\frac{2}{5\cdot 11\cdot 49} = \frac{4}{11\cdot 49} = 0.00742115$$

(2) If three different people win, then this is the disjoint union of three out of ten people winning one ticket. But for each of these three people, only one out of ten tickets are chosen. So we have

$$\binom{10}{3}\cdot\frac{\binom{10}{1}\binom{10}{1}\binom{10}{1}}{\binom{100}{3}} = 10\cdot 10\cdot 10\cdot\frac{2}{5\cdot 11\cdot 49} = \frac{100\cdot 4}{11\cdot 49} = 0.742115$$

(3) Finally, we know that if two different people win that means neither three people won nor exactly one person won. Since exactly three tickets are handed out, we end up with:

$$1 - 0.00742115 - 0.742115 = 0.250464$$

Alternatively, we could have seen a pattern above and done the following:

$$2\cdot\binom{10}{2}\cdot\frac{\binom{10}{2}\binom{10}{1}}{\binom{100}{3}}$$

to get the right answer.

□

Exercise 47 Let X be the number of tails in three random tosses of a fair coin. Write the distribution of X in a table and find the distribution of $|X-1|$.
Answer. We know that X should have the binomial distribution which is given by

$$P(X=x) = \binom{3}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x}$$

Plugging in, 0, 1, 2 and 3 for x we get:

x	0	1	2	3
$P(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

If we're trying to find $|X-1|$ then we know:

$$|0-1| = 1 \qquad |1-1| = 0 \qquad |2-1| = 1 \qquad |3-1| = 2$$

Therefore

x	0	1	2
$P(\lvert X-1\rvert = x)$	$\frac{3}{8}$	$\frac{1}{8}+\frac{3}{8}=\frac{1}{2}$	$\frac{1}{8}$

□

Exercise 48 Let X and Y be the numbers obtained in two pullings of tickets out of a hat which has four tickets labelled 1, 2, 3, 4. Find the joint distribution table for X and Y if we do the sampling with and without replacement. In each of the cases find the probability $P(X \leq Y)$.
Answer. If we do it with replacement we get:

	$X=1$	$X=2$	$X=3$	$X=4$	Total
$Y=1$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$Y=2$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$Y=3$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
$Y=4$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$
Total	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

To find $P(X \leq Y)$ we just need to look at the bottom left triangle of the 4×4 grid. Since there are 10 entries, each equal to $\frac{1}{16}$ we get:

$$P(X \leq Y) = 10 \cdot \frac{1}{16} = \frac{5}{8}$$

If we do it without replacement

	$X=1$	$X=2$	$X=3$	$X=4$	Total
$Y=1$	0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
$Y=2$	$\frac{1}{12}$	0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$
$Y=3$	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{12}$	$\frac{1}{4}$
$Y=4$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0	$\frac{1}{4}$
Total	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

Just as before, we look at the bottom left triangle to determine $P(X \leq Y)$.

$$P(X \leq Y) = 4 \cdot 0 + 6 \cdot \frac{1}{12} = \frac{1}{2}$$

□

Exercise 49 Let X_1 and X_2 be the numbers obtained by rolling two rolls of a fair die. Let $Y_1 = \max(X_1, X_2)$ and $Y_2 = \min(X_1, X_2)$. We already calculated the joint distribution table for X_1 and X_2 in class. What is the joint distribution table for Y_1 and Y_2?
Answer. We can use the probability distribution from class to determine the desired probability distribution:

$X_2 \backslash X_1$	1	2	3	4	5	6	Total
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
Total	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Using the table above, we can calculate the distribution for Y_1 and Y_2

$Y_2 \backslash Y_1$	1	2	3	4	5	6	Total
1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{11}{36}$
2	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
3	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{7}{36}$
4	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{5}{36}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$
6	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
Total	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	1

For example, in the second row and the fourth column we have $\frac{2}{36}$ since there are two ways for the minimum to be 2 and the maximum to be 4: $(2,4)$ and $(4,2)$. In the 5th row and second column we have 0 since there are no ways for the minimum to be 5 and the maximum to be 2 of any two numbers. □

Exercise 50 Let X be the number of heads which appear in 20 random tosses of a fair coin. Let Y be a number which is randomly picked from a hat with the numbers $\{0,1,2,\dots,20\}$ (independently from X). Let $Z = \max(X,Y)$. Find a formula for $P(Z=k)$ where $0 \le k \le 20$.
Answer. We've seen this method a couple times now, but we can find $P(Z=k)$ by:

$$\begin{aligned} P(Z=k) &= P(X=k, Y<k) + P(X<k, Y=k) + P(X=k, Y=k) \\ &= \binom{20}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{20-k}\cdot\frac{k}{21} + \frac{1}{21}\sum_{i=0}^{k-1}\binom{20}{i}\left(\frac{1}{2}\right)^i\left(\frac{1}{2}\right)^{20-i} + \binom{20}{k}\frac{1}{2^{20}}\frac{1}{21} \\ &= \binom{20}{k}\frac{k}{21\cdot 2^{20}} + \sum_{i=0}^{k}\binom{20}{i}\frac{1}{21\cdot 2^{20}} \end{aligned}$$

□

Exercise 51 Suppose that we have two dice which are not fair. The number i appears with probability p_i for the first die and with probability p_i' for the second die. Let S be the sum of numbers rolled with these two dice. Find the formula for $P(S=2)$, $P(S=7)$ and $P(S=12)$.
Answer. Since these are independent dice rolls, we know that

$$P(S=k) = \sum_{i+j=k} p_i p_j'$$

The best way to find this number is to think (or draw out) the joint distribution table and think of which rows and columns you want. We get:

$$\begin{gathered} P(S=2) = p_1 p_1' \\ P(S=12) = p_6 p_6' \\ P(S=7) = P_1 P_6' + p_2 p_5' + p_3 p_4' + p_4 p_3' + p_5 p_2' + p_6 p_1' \end{gathered}$$

□