STAT 630 Problem Set 2

Nikko Cleri

September 17, 2021

Question I.

(a) Prove Bonferroni's inequality $P(A \cap B) \geq P(A) + P(B) - 1$

We can start with the interpretation that

$$\begin{aligned} P((A \cap B)^c) &= 1 - P(A \cap B) \\ &= P(A^c \cup B^c) \end{aligned}$$

We know from that for the union of two events $P(A \cup B) \leq P(A) + P(B)$, which gives:

$$\begin{aligned} P(A^c \cup B^c) &\leq P(A^c) + P(B^c) \\ &\leq 1 - P(A) + 1 - P(B) \end{aligned}$$

combining this with the first line of the proof we have

$$\begin{aligned} 1 - P(A \cap B) &\leq P(A^c) + P(B^c) \\ &\leq 1 - P(A) + 1 - P(B) \\ P(A \cap B) &\geq P(A) + P(B) - 1 \quad \square \end{aligned}$$

(b) In statistics we often talk about the event that our statistical procedure will lead to a correct (true) conclusion. Suppose A and B are such events (for two different procedures, but in the same experiment) and each has probability 95%. According to Bonferroni's inequality, what can we say about the chance that both procedures lead to correct conclusions?

The probability that both events A and B occur ($P(A \cap B)$) is at least $.95 + .95 - 1 = .9$.

(c) Extrapolate to 3 events. Specifically, suppose 3 statistical procedures each have probability of $1 - \alpha/3$ of resulting in a correct conclusion and show that the probability that all 3 are correct is at least 1-α.

For 3 events, Bonferroni's inequality becomes:

$$P(A \cap B \cap C) \geq P(A) + P(B) + P(C) - 2$$

So we have

$$\begin{aligned} P(A \cap B \cap C) &\geq 1 - \alpha/3 + 1 - \alpha/3 + 1 - \alpha/3 - 2 \\ &\geq 3 - \alpha - 2 \\ &\geq 1 - \alpha \end{aligned}$$

Question II.

1.5.9

Suppose we roll two fair six-sided dice, one red and one blue. Let A be the event that the two dice show the same value. Let B be the event that the sum of the two dice is equal to 12. Let C be the event that the red die shows 4. Let D be the event that the blue die shows 4.

(a) Are A and B independent?

$P(A) = 1/6$ and $P(B) = 1/36$. The intersection (rolling two sixes) has a probability $P(A \cap B) = P(B) = 1/36 \neq (1/6)(1/36)$, so A and B are not independent.

(b) Are A and C independent?

$P(A) = 1/6$ and $P(C) = 1/6$. The intersection (rolling two fours) has a probability $P(A \cap C) = 1/36 = (1/6)(1/6)$, so A and C are independent.

(c) Are A and D independent?

Identically to the previous part, $P(A) = 1/6$ and $P(D) = 1/6$. The intersection (rolling two fours) has a probability $P(A \cap D) = 1/36 = (1/6)(1/6)$, so A and D are independent.

(d) Are C and D independent?

Here we have $P(C) = 1/6$ and $P(D) = 1/6$. The intersection (rolling two fours) has a probability $P(C \cap D) = 1/36 = (1/6)(1/6)$, so C and D are independent.

(e) Are A, C, and D all independent?

Again we have $P(A) = 1/6$, $P(C) = 1/6$, and $P(D) = 1/6$ The intersect of all three is again rolling two fours, $P(A \cap C \cap D) = 1/36$, but this is not equal to the product of the three events $P(A)P(C)P(D) = 1/216 \neq 1/36$, so A, C, and D are not all independent.

1.5.14

Prove that A and B are independent if and only if A^c and B are independent.

In one direction, assume A and B are independent, so $P(A \cap B) = P(A)P(B)$. We need to prove:

$$P(A^c \cap B) = P(A^c)P(B)$$

Beginning with the right hand side:

$$\begin{aligned} P(A^c)P(B) &= (1 - P(A))P(B) \\ &= P(B) - P(A)P(B) \\ &= P(B) - P(A \cap B) \\ &= P(B \cap A^c) \end{aligned}$$

proves the lemma one in one direction. In the other direction, assume A^c and B are independent, so $P(A^c \cap B) = P(A^c)P(B)$, and prove

$$P(A \cap B) = P(A)P(B)$$

Start with the right hand side:

$$\begin{aligned}P(A)P(B) &= (1 - P(A^c))P(B) \\ &= P(B) - P(A^c)P(B) \\ &= P(B) - P(A^c \cap B) \\ &= P(A \cap B)\end{aligned}$$

Which proves the lemma in the other direction.

1.5.18(a,b,c)

(The Monty Hall problem) Suppose there are three doors, labeled A, B, and C. A new car is behind one of the three doors, but you don't know which. You select one of the doors, say, door A. The host then opens one of doors B or C, as follows: If the car is behind B, then they open C; if the car is behind C, then they open B; if the car is behind A, then they open either B or C with probability $1/2$ each. (In any case, the door opened by the host will not have the car behind it.) The host then gives you the option of either sticking with your original door choice (i.e., A), or switching to the remaining unopened door (i.e., whichever of B or C the host did not open). You then win (i.e., get to keep the car) if and only if the car is behind your final door selection. (Source: Parade Magazine, "Ask Marilyn" column, September 9, 1990.) Suppose for definiteness that the host opens door B.

(a) If you stick with your original choice (i.e., door A), conditional on the host having opened door B, then what is your probability of winning? (Hint: First condition on the true location of the car. Then use Theorem 1.5.2.)

From Bayes' Theorem we know

$$P(A|B) = \frac{P(A)}{P(B)}P(B|A)$$

where $P(A|B)$ is the probability that door A hides the car given door B was opened. $P(A) = 1/3$ is the probability that door A hides the car, $P(B) = 1/2$ is the probability that door B is opened if door A is the initial selection, and $P(B|A) = 1/2$ is the probability that door B is opened if door A hides the car. This gives

$$\begin{aligned}P(A|B) &= \frac{P(A)}{P(B)}P(B|A) \\ &= \frac{1/3}{1/2}1/2 \\ &= \frac{1}{3}\end{aligned}$$

The probability of winning without switching is $1/3$.

(b) If you switch to the remaining door (i.e., door C), conditional on the host having opened door B, then what is your probability of winning?

This answer is simply the complement of the answer in part (a), so the probability of winning after switching is $2/3$.

(c) Do you find the result of parts (a) and (b) surprising? How could you design a physical experiment to verify the result?

This is a well known exercise in probability, however the result does not feel intuitive. To experimentally support these probabilities, we can perform n (some large number) of trials where we stay with our selection, and track our success rate. We then can repeat this for another n trials switching our selection every time, and again record our success rate. This should yield the probabilities from (a) and (b) in the large-n limit.

Question III.

The system shown below has five components which act independently. Each component fails with probability p. Find the probability that the system fails. See Figure 1.

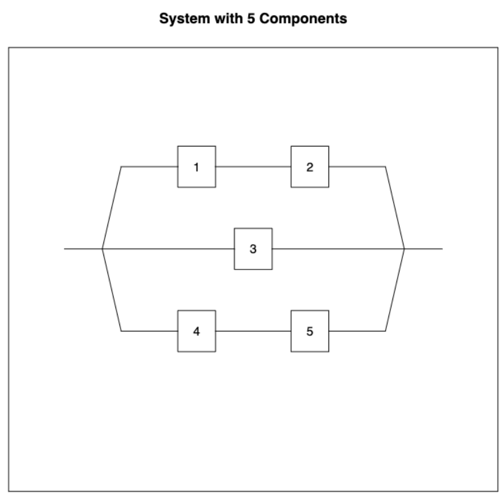

Figure 1: System for Question 3.

In order for the entire system to fail, all three routes of the system must fail simultaneously. The probability of success for any one component is $1 - p$.

The simplest route is the middle, with component 3, which fails with a probability p.

For the first route, with components 1 and 2, the probability of failure is one minus the probability of both components working, which will be $1 - (1 - p)^2$. The third route, with components 4 and 5, similarly fails with a probability $1 - (1 - p)^2$.

For the entire system, we multiply the failure probabilities together, to find that the system fails with a probability $p\left[1 - (1 - p)^2\right]^2$.

Question IV.

If a parent has genotype Aa, he transmits either A or a to an offspring, each with probability 1/2. The gene he transmits to one offspring is independent of the gene he transmits to all other offspring. Consider a parent with three children (labeled 1,2,3) and the following events: B=\{1 and 2 have the same gene}, C={2 and 3 have the same gene}, D={1 and 3 have the same gene}. Show that all these events are pairwise independent, but not mutually independent.

Testing for pairwise independence:

Between B and C:

$$\begin{aligned} P(B \cap C) &= P(\text{all same}) \\ &= \frac{2}{8} = \frac{1}{4} \\ P(B)P(C) &= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \\ P(B \cap C) &= P(B)P(C) \end{aligned}$$

So B and C are independent. The other two pairs follow similarly:

$$\begin{aligned} P(B \cap D) &= P(\text{all same}) \\ &= \frac{2}{8} = \frac{1}{4} \\ P(B)P(D) &= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \\ P(B \cap D) &= P(B)P(D) \end{aligned}$$

$$\begin{aligned} P(C \cap D) &= P(\text{all same}) \\ &= \frac{2}{8} = \frac{1}{4} \\ P(C)P(D) &= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \\ P(C \cap D) &= P(C)P(D) \end{aligned}$$

So B, C, and D are pairwise independent. Finally, to show they are not mutually independent, we have:

$$\begin{aligned} P(B \cap C \cap D) &= P(\text{all same}) \\ &= \frac{2}{8} = \frac{1}{4} \\ P(B)P(C)P(D) &= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8} \\ P(B \cap C \cap D) &\neq P(B)P(C)P(D) \end{aligned}$$

Question V.

2.1.5

(a) Let A and B be events, and let $X = I_A I_B$. Is X an indicator function? If yes, then of what event?

X is an indicator of the intersection $A \cap B$.

(b) Show that $I_{A\cup B} = \max(I_A, I_B)$.

If the indicator functions are defined such that

$$I_J(s) = \begin{cases} 1 & s \in J \\ 0 & s \notin J \end{cases}$$

for some random variable J, then we have

$$I_A(s) = \begin{cases} 1 & s \in A \\ 0 & s \notin A \end{cases}$$
$$I_B(s) = \begin{cases} 1 & s \in B \\ 0 & s \notin B \end{cases}$$
$$I_{A\cup B}(s) = \begin{cases} 1 & s \in A \cup B \\ 0 & s \notin A \cup B \end{cases}$$

So $I_{A\cup B}(s)$ takes the maximum value of I_A and I_B.

Question VI.

2.1.8

Compute $W(s)$ and $Z(s)$ for all $s \in S$

Let $S = \{1, 2, 3, 4, 5\}$, $X = I_{\{1,2,3\}}$, $Y = I_{\{2,3\}}$, and $Z = I_{\{3,4,5\}}$. Let $W = X - Y + Z$.

$$W(1) = 1 - 0 + 0 = 1$$
$$W(2) = 1 - 1 + 0 = 0$$
$$W(3) = 1 - 1 + 1 = 1$$
$$W(4) = 0 - 0 + 1 = 1$$
$$W(5) = 0 - 0 + 1 = 1$$

$$Z(1) = 0$$
$$Z(2) = 0$$
$$Z(3) = 1$$
$$Z(4) = 1$$
$$Z(5) = 1$$

So $W \geq Z \ \forall \ s \in S$

Question VII.

2.2.4

Suppose we roll one fair six-sided die, and let Z be the number showing. Let $W = Z^3 + 4$ and let $V = \sqrt{Z}$.

(a) Compute $P(W = w)$ for every real number w.

$P(W = w) = 0$ for any $w \notin \{5, 12, 31, 68, 129, 220\}$. Each of these six values of W is equally probably with probability $P(W = w) = 1/6$.

(b) Compute $P(V = v)$ for every real number v.

$P(V = v) = 0$ for any $v \notin \{1, \sqrt{2}, \sqrt{3}, 2, \sqrt{5}, \sqrt{6}\}$. Each of these six values of V is equally probably with probability $P(V = v) = 1/6$.

(c) Compute $P(ZW = x)$ for every real number x.

$P(ZW = x) = 0$ for any $x \notin \{5, 24, 93, 272, 645, 1320\}$. Each of these six values of ZW is equally probably with probability $P(ZW = x) = 1/6$.

(d) Compute $P(VW = y)$ for every real number y.

$P(VW = y) = 0$ for any $y \notin \{5, 12\sqrt{2}, 31\sqrt{3}, 136, 129\sqrt{5}, 220\sqrt{6}\}$. Each of these six values of VW is equally probably with probability $P(VW = y) = 1/6$.

(e) Compute $P(V + W = r)$ for every real number r.

$P(V + W = r) = 0$ for any $r \notin \{6, 12 + \sqrt{2}, 31 + \sqrt{3}, 70, 129 + \sqrt{5}, 220 + \sqrt{6}\}$. Each of these six values of $V + W$ is equally probably with probability $P(V + W = r) = 1/6$.

Question VIII.

2.3.9

Let $Z \sim$Negative Binomial$(3, 1/4)$. Compute $P(Z \le 2)$.

$$
\begin{aligned}
p_Z(z) &= \binom{r-1+z}{r-1}\theta^r(1-\theta)^z \quad z = 0, 1... \\
P(Z \le 2) &= P(Z = 0) + P(Z = 1) + P(Z = 2) \\
&= \frac{1}{64} + \frac{9}{256} + \frac{27}{512} \\
&= \frac{53}{512}
\end{aligned}
$$

2.3.10

Let $X \sim$Geometric$(3, 1/4)$. Compute $P(X^2 \le 15)$.

$$
\begin{aligned}
p_X(x) &= \theta(1-\theta)^x \quad x = 0, 1, 2... \\
P(X^2 \leq 15) &= P(X \leq \sqrt{15}) \\
&= P(X=0) + P(X=1) + P(X=2) + P(X=3) \\
&= \frac{1}{4} + \frac{3}{16} + \frac{9}{64} + \frac{27}{256} \\
&= \frac{176}{256}
\end{aligned}
$$

2.3.13

Let X ~Hypergeometric$(20, 7, 8)$. What is the probability that $X = 3$? What is the probability that $X = 8$?

$$
\begin{aligned}
p_X(x) &= \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \quad x = M_1...M_2 \\
P(X=3) &= \frac{231}{646} \\
P(X=8) &= 0
\end{aligned}
$$

2.3.14

Suppose that a symmetrical die is rolled 20 independent times, and each time we record whether or not the event $\{2, 3, 5, 6\}$ has occurred.

(a) What is the distribution of the number of times this event occurs in 20 rolls?

This is a binomial where success is the event $\{2, 3, 5, 6\}$ occurring.

$$
\begin{aligned}
p_X(x) &= \binom{20}{x}\left(\frac{4}{6}\right)^x\left(1-\frac{4}{6}\right)^{20-x} \quad x = 0, 1...20 \\
&= 0 \quad \text{otherwise}
\end{aligned}
$$

(b) Calculate the probability that the event occurs five times.

$$
\begin{aligned}
p_X(5) &= \binom{20}{5}\left(\frac{4}{6}\right)^5\left(1-\frac{4}{6}\right)^{20-5} \\
&= 0.00014
\end{aligned}
$$

Question IX.

2.3.15 For (b) use the interpretation that the first basket is obtained on the 10th attempt, and for (c) that the second basket is obtained on the 10th attempt.

Suppose that a basketball player sinks a basket from a certain position on the court with probability 0.35.

(a) What is the probability that the player sinks three baskets in 10 independent throws?

This is a binomial with $\theta = .35$. We have

$$\begin{aligned} p_X(3) &= \binom{10}{3} .35^3 \, (1 - .35)^{10-3} \\ &= 0.252 \end{aligned}$$

(b) What is the probability that the player throws 10 times before obtaining the first basket?

This is a geometric with $\theta = .35$. (The player misses 9 times then makes the tenth).

$$\begin{aligned} p_X(9) &= \theta^1 (1 - \theta)^9 \\ &= 0.007 \end{aligned}$$

(c) What is the probability that the player throws 10 times before obtaining two baskets?

Now we have a negative binomial with $\theta = .35$ and $r = 2$.

$$\begin{aligned} p_Z(9) &= \binom{2 - 1 + 9}{2 - 1} \theta^2 (1 - \theta)^8 \\ &= 0.039 \end{aligned}$$

Question X.

Use `R` to check the accuracy of the binomial approximation to the hypergeometric distribution. You can use the `dbinom` and `dhyper` functions to compute probabilities for the two distributions. Compare the distributions for $n = 10$, $M/N = 0.6$ and $N = 50, \ 100, \ 1000$. The syntax in `R` for computing the hypergeometric probability mass function is `dhyper(x,M,N-M,n)`. Note that x can be a vector in either function. So, for example, `dbinom(0:3,10,.4)` returns a vector with binomial(10,.4) probabilities for $x = 0, 1, 2, 3$.

```
> dbinom(0:10,10,.6)
 [1] 0.0001048576 0.0015728640 0.0106168320 0.0424673280 0.1114767360 0.2006581248 0.2508226560
 [8] 0.2149908480 0.1209323520 0.0403107840 0.0060466176
> dhyper(0:10, .6*1000, 1000-.6*1000, 10)
 [1] 9.793903e-05 1.502901e-03 1.033436e-02 4.193348e-02 1.111929e-01 2.013295e-01 2.520856e-01
 [8] 2.155291e-01 1.204228e-01 3.970499e-02 5.866412e-03
> dhyper(0:10, .6*100, 100-.6*100, 10)
 [1] 4.896854e-05 9.477781e-04 7.863597e-03 3.685565e-02 1.081280e-01 2.076057e-01 2.643128e-01
 [8] 2.204307e-01 1.152911e-01 3.416032e-02 4.355441e-03
> dhyper(0:10, .6*50, 50-.6*50, 10)
 [1] 1.798588e-05 4.905241e-04 5.334450e-03 3.063889e-02 1.034063e-01 2.150850e-01 2.800586e-01
 [8] 2.259296e-01 1.082579e-01 2.785585e-02 2.924864e-03
```

Figure 2: R results for Question 10. The hypergeometric and binomial agree more strongly at large N. This is visible by looking at the difference between the `dbinom` and `dhyper` outputs at a given x, where the differences between the $N = 1000$ hypergeometric and binomial are the smallest.