Stat 512
Homework key 1

September 30, 2015

REGULAR PROBLEMS

1. Let A, B, and C be measurable events for some probability measure P on sample space Ω. Further suppose $P(A) > 0, P(B) > 0$ and $P(C) > 0$, with none of the pairs of events mutually exclusive.

   (a) Show $P(AB|C) = P(A|BC).P(B|C)$.

   (b) Show $P(A|B).P(B) = P(B|A).P(A)$.

   (c) Show $P(ABC) = P(A|BC).P(B|C).P(C)$.

   (d) What general condition is sufficient for $P(ABC) = P(A|BC).P(B|AC).P(C|AB)$ to be true? Is that condition necessary?

   Solution.

   (a)

$$
\begin{aligned}
P(AB|C) =& \frac{P(ABC)}{P(C)} \\
=& P(A|BC)\frac{P(BC)}{P(C)} \\
=& P(A|BC).P(B|C).
\end{aligned}
$$

   (b)

$$
\begin{aligned}
& P(AB) = P(A|B).P(B);\ P(AB) = P(B|A).P(A). \\
\therefore\ & P(A|B).P(B) = P(B|A).P(A).
\end{aligned}
$$

   (c)

$$
\begin{aligned}
P(ABC) =& P(A|BC)P(BC) \\
=& P(A|BC).P(B|C).P(C).
\end{aligned}
$$

   (d) If A, B, C are independent events, then $P(A|BC) = P(A), P(B|AC) = P(B), P(C|AB) = P(C)$. Also, $P(ABC) = P(A)P(B)P(C)$. Hence, $P(ABC) = P(A|BC).P(B|AC).P(C|AB)$. Hence, a sufficient condition is the independence of the events A, B, C.

   We will provide a counterexample to establish that the above-mentioned condition is not necessary. Suppose A and B are independent events. C is such an event that $A \subset C, B \subset C$ and $P(C) < 1$ (Notice that $C = A \cup B$ will work). Therefore, $A \cap C = A$ and $B \cap C = B$. Then $AB \subset C$, which leads to $P(ABC) = P(AB) = P(A)P(B)$. Now since $A \cap C = A$ and $B \cap C = B$,

$$P(A|BC) = P(A|B) = P(A)$$

   as A and B are independent by assumption. Similarly,

$$P(B|AC) = P(B|A) = P(B).$$

Now,

$$P(C|AB) = \frac{P(ABC)}{P(AB)} = \frac{P(AB)}{P(AB)} = 1.$$

So,

$$P(A|BC)P(B|AC)P(C|AB) = P(A)P(B) = P(AB) = P(ABC).$$

We still need to check that A, B, C are not independent.

$$P(AC) = P(A) > P(A)P(C)$$

since $P(C) < 1$. Hence A and C are not independent.

2. Let A and B be independent measurable events for some probability measure P on sample space Ω.

   (a) Show A and B^c are independent measurable events.

   (b) Show A^c and B are independent measurable events.

   (c) Show A^c and B^c are independent measurable events.

Solution.

(a) Notice that $AB^C = A \cap (AB)^C$

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

Hence, proved.

(b)

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

Hence, proved.

(c) Using (a) we get that A and B^C are independent. Now by (b), if A and B are independent, A^C and B are independent. Combining (a) and (b) we get that, A^C and B^C are independent.

3. Let $A_1, A_2, \ldots, A_n$ be measurable events for some probability measure P on sample space Ω.

   (a) Show $P(\cap_{i=1}^n A_i^C) = P([\cup_{i=1}^n A_i]^C)$.

   (b) Show $P(\cup_{i=1}^n A_i^C) = P([\cap_{i=1}^n A_i]^C)$.

Solution.

(a) Let $U = \cap_{i=1}^{n} A_i^C$ and $V = [\cup_{i=1}^{n} A_i]^C$. We will show that $U \subseteq V$ and $V \subseteq U$, hence $U = V$. This will lead to

$$P(\cap_{i=1}^{n} A_i^C) = P([\cup_{i=1}^{n} A_i]^C).$$

Now let $\omega \in U$. Hence, $\omega \in A_i^C$ for all i. Therefore, $\omega \notin A_i$ for all i. This can be written as $\omega \notin \cup A_i$ or $\omega \in [\cup_{i=1}^{n} A_i]^C$.. Hence $U \subseteq V$.

Similarly suppose,

$$\begin{aligned}
&\omega \in V. \\
&\Rightarrow \omega \notin A_i \ \forall i. \\
&\Rightarrow \omega \in A_i^C \ \forall i. \\
&\Rightarrow \omega \in \cap_{i=1}^{n} A_i^C = U. \\
&\Rightarrow V \subseteq U.
\end{aligned}$$

This completes the proof.

(b) Let $U = \cup_{i=1}^{n} A_i^C$ and $V = [\cap_{i=1}^{n} A_i]^C$. We will show that $U \subseteq V$ and $V \subseteq U$, hence $U = V$. This will lead to

$$P(\cup_{i=1}^{n} A_i^C) = P([\cap_{i=1}^{n} A_i]^C).$$

Now, let $\omega \in U$. Hence, $\exists\ i$ such that $\omega \notin A_i$. Hence, $\omega \notin \cap_{i=1}^{n} A_i$. Therefore, $\omega \in V$. Hence, $U \subseteq V$.

Let $\omega \in V$. Hence, $\exists\ i$ such that $\omega \notin A_i$. Hence, $\omega \in \cup_{i=1}^{n} A_i^C$. Therefore, $V \subseteq U$ which completes the proof.

4. Let A and B be measurable events for some probability measure P on sample space Ω.

(a) Assume $P(A) > 0$ and $P(B) > 0$. Show that if A and B are mutually exclusive events, they are not independent.

(b) Assume $P(A) > 0$ and $P(B) > 0$. Show that if A and B are independent events, they are not mutually exclusive.

(c) Find necessary and sufficient conditions on A and B that they would be both mutually exclusive and independent.

Solution.

(a) If A and B are mutually exclusive (disjoint) events, then

$$P(A|B) = 0.$$

If A and B are independent events we have further that

$$P(A|B) = P(A).$$

But $P(A) > 0$ by hypothesis, so $P(A|B) \neq P(A)$. Hence, A and B can not be independent events.

(b) If A and B are independent events, then

$$P(A|B) = P(A) > 0.$$

Hence, $P(A|B) \neq 0$ so that A and B can not be mutually exclusive.

(Note: the above arguments remain valid if the roles of B and A are reversed.)

(c) If A and B are mutually exclusive events and independent, then $P(A|B) = P(B|A) = 0$ and $P(A|B) = P(A)$, $P(B|A) = P(B)$. Hence, $P(A) = P(B) = 0$ is necessary. If A and B are disjoint sets with $P(A) = P(B) = 0$, then $P(AB) = 0 = P(A)P(B)$, so A and B are also independent. Hence, a necessary and sufficient condition for A and B to be both mutually exclusive and independent is for them to be disjoint sets each having probability zero. Note that this situation is degenerate, since conditional probabilities are typically defined only in the case that the set being conditioned on has positive probability.

5. Let A and B be measurable events for some probability measure P on sample space Ω. Further suppose $P(A) = 0.3$, $P(A \cup B) = 0.8$.

(a) Find $P(B)$ if A and B are mutually exclusive events.

(b) Find $P(B)$ if A and B are independent events.

(c) Find $P(B)$ if $P(A|B) = 0.3$.

(d) Can $P(A|B) = 0.5$? If so, find $P(B)$. If not, explain why not.

Solution.

(a) If A and B are disjoint (mutually exclusive), then $P(AB) = 0$ and

$$P(A \cup B) = P(A) + P(B) - P(AB) = 0.3 + P(B) = 0.8,$$

which is solved by $P(B) = 0.5$.

(b) If A and B are independent, $P(AB) = P(A)P(B) = 0.3P(B)$, so

$$P(A \cup B) = 0.3 + P(B) - 0.3P(B) = 0.3 + 0.7P(B) = 0.8$$

and, solving for $P(B)$, we obtain $P(B) = 5/7 \approx 0.7143$.

(c) If $P(A|B) = 0.3$, then

$$P(AB) = P(A|B)P(B) = 0.3P(B).$$

As in (b), this leads to the equation

$$P(A \cup B) = 0.3 + P(B) - 0.3P(B) = 0.3 + 0.7P(B) = 0.8$$

and, solving for $P(B)$, we obtain $P(B) = 5/7 \approx 0.7143$.

(d) If $P(A|B) = 0.5$, then $P(AB) = P(A|B)P(B) = 0.5P(B)$ and we have that

$$P(A \cup B) = 0.3 + P(B) - 0.5P(B) = 0.3 + 0.5P(B) = 0.8,$$

implying that $P(B) = 1$. If this were true, then we obtain the contradiction $P(A \cup B) = 1 \neq 0.8$. Hence, $P(A|B)$ cannot equal 0.5.

MORE INVOLVED PROBLEMS

6. Consider an experiment of n independent rolls of a fair die. Let X_k be the random variable measuring the number showing on top of the die on the kth roll. Hence, the sample space for X_k is $\Omega_X = \{1, 2, 3, 4, 5, 6\}$, with $P_X(X_k = \omega) = \frac{1}{6}$ for all $\omega \in \Omega_X$ and for all $k = 1, \ldots, n$.

   Define $Y_k = \min(X_1, \ldots, X_n)$ and $W_k = \max(X_1, \ldots, X_n)$ for $k = 1, \ldots, n$.

   (a) Provide formulas for the cumulative distribution functions $F_{Xk}(x) = Pr(X_k \leq x)$, $F_{Yk}(y) = Pr(Y_k \leq y)$, and $F_{Wk}(w) = Pr(W_k \leq w)$.

   (b) Provide plots of F_{Xk}, F_{Yk}, and F_{Wk} for $k = 1, 2$, and 5.

   (c) Comment on how this problem might be indicative of the "multiple comparison problem" in which the smallest of multiple p-values is used to define an association.

   Solution.

   (a) Let $h(x) = 1[x \leq c]$ be the indicator function equal to 1 if $x \leq c$ and 0 otherwise. Then,

$$\begin{aligned}
F_{Xk}(x) &= Pr(X_k \leq x) \\
&= \frac{1}{6}\sum_{i=1}^{n} 1[x \leq i] \\
&= \begin{cases} 0 & x < 1 \\ \frac{\lfloor x \rfloor}{6} & 1 \leq x < 6 \\ 1 & 6 \leq x \end{cases}
\end{aligned}$$

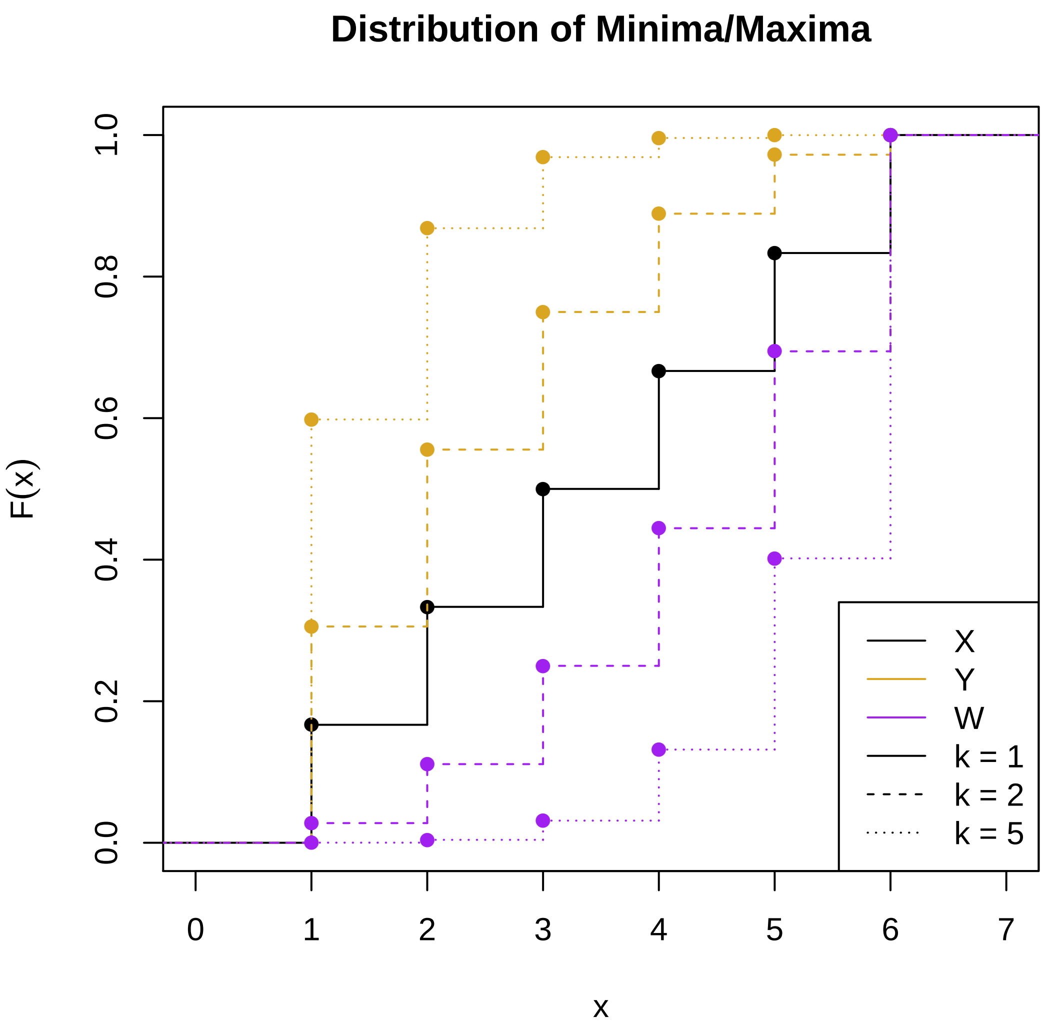
   To calculate the distribution of Y_k, we note that all of the X_k must be at least as large as their minimum. Then, we use independence of the die rolls, as follows:

$$\begin{aligned}
F_{Yk}(y) &= Pr(Y_k \leq y) \\
&= Pr(\min(X_1, \ldots, X_k) \leq y) \\
&= 1 - Pr(\min(X_1, \ldots, X_k) > y) \\
&= 1 - Pr(X_1 > y, \ldots, X_k > y) \\
&= 1 - \prod_{i=1}^{k} Pr(X_i > y) \\
&= 1 - Pr(X_k > y)^k \\
&= 1 - (1 - Pr(X_k \leq y))^k \\
&= 1 - (1 - F_{Xk}(y))^k \\
&= \begin{cases} 0 & y < 1 \\ 1 - (1 - \lfloor y \rfloor / 6)^k & 1 \leq y < 6 \\ 1 & 6 \leq y \end{cases}
\end{aligned}$$

   Similarly, the distribution of W_k follows from recognizing that all the X_k are bounded above by their maximum and again using their independence:

$$\begin{aligned} F_{Wk}(w) &= Pr(W_k \le w) \\ &= Pr(\max(X_1, \ldots, X_k) \le w) \\ &= Pr(X_1 \le w, \ldots, X_k \le w) \\ &= \prod_{i=1}^{k} Pr(X_i \le w) \\ &= Pr(X_k \le w)^k \\ &= F_{Xk}(w)^k \\ &= \begin{cases} 0 & w < 1 \\ (\lfloor w \rfloor / 6)^k & 1 \le w < 6 \\ 1 & 6 \le w \end{cases} \end{aligned}$$

(b) Note that in the case that $k = 1$, that X_k, Y_k, and W_k are equal in distribution. The distribution of the individual rolls does not depend on k. However, from the plot below, we see that as k increases, Y_k has more and more of its mass shifted towards 1 and W_k sees most of its mass shift towards 6.



(c) If we are interested in testing the statistical hypothesis $H_0^{(1)}$ vs. $H_1^{(1)}$ with a false positive rate of 1/6, we could simply roll a 6-sided die, rejecting $H_0^{(1)}$ in favor of $H_1^{(1)}$ if we roll a 1, and not rejecting if we roll any greater number.

Were we to repeat this experiment for k pairs of hypotheses, $H_0^{(1)}$ vs. $H_1^{(1)}$, ..., and $H_0^{(k)}$ vs. $H_1^{(k)}$, the probability that our minimum roll is 1 (equivalently, that we make one *or more* false rejections) is $1-(1-1/6)^k = 1-(5/6)^k$, which tends towards 1 as k increases. This means we are making false discoveries far more often than 1/6-th of the time under this decision rule.

It can be shown in general that if the probability of making a false rejection is α for one test, testing k hypotheses at this same level results in a probability of *at least one* false rejection of $1-(1-\alpha)^k$.

7. Let A, B and C be measurable events for some probability measure P on sample space. Further suppose $P(A) > 0$, $P(B) > 0$, and $P(C) > 0$, with none of the pairs of events mutually exclusive. Simpson's paradox states that it is possible to have

$$P(A|BC) > P(A|B^cC),$$

$$P(A|BC^c) > P(A|B^cC^c)$$

but

$$P(A|B) < P(A|B^c).$$

(a) Illustrate how Simpson's paradox might lead to erroneous conclusions when investigating smoking effects on lung cancer deaths across two different countries. Consider analyses that merely compare smokers to nonsmokers and finds that smoking is protective against lung cancer death versus an analysis that compares smokers to nonsmokers within each country. Explicitly provide probabilities for each event that would lead to such a setting. (It is sufficient to consider dichotomized smoking behavior.)

(b) Prove the following (including showing that the conditions are not necessary). For events A, B, and C, with

$$P(A|BC) > P(A|B^cC),$$

$$P(A|BC^c) > P(A|B^cC^c)$$

then either of the following conditions

- B and C are independent, OR
- A and C are independent when conditioned on B (so $P(AC|B) = P(A|B)P(C|B)$)

are sufficient (but not necessary) to guarantee

$$P(A|B) > P(A|B^c).$$

(a) Consider the following scenario. Suppose for country 1 (we will denote by C) we get the following counts.

	Smoker	Non-smoker
Death by lung cancer	30	290
No lung cancer	70	710

Let A denote the event of death by lung cancer and B denote the event of smoking. Then for country 1 i.e. event C, we get,

$$P(A|BC) = .3;\ P(A|B^cC) = .29.$$

Suppose for country 2 (we will denote by C^c) we get the following counts.

$$P(A|BC) > P(A|B^cC),$$
$$P(A|BC^c) > P(A|B^cC^c)$$

	Smoker	Non-smoker
Death by lung cancer	20	10
No lung cancer	80	90

Then for country 2 i.e. event C^c, we get,

$$P(A|BC^c) = .2;\ P(A|B^cC^c) = .1.$$

Therefore, we see that for both countries, death by lung cancer is observed more frequently in smokers since

Combining the two tables we see that a study which does not consider the stratification by countries will obtain the following table:

	Smoker	Non-smoker
Death by lung cancer	50	300
No lung cancer	150	800

$$P(A|B) = .333\ldots;\ P(A|B^c) = .375.$$
$$P(A|B) < P(A|B^c)$$

Hence, this study will deduce that smoking protects from lung cancer.

(b) • Suppose B and C are independent. Using 1.(a) we get

$$P(AC|B) = P(A|BC)P(C|B) = P(A|BC)P(C) \tag{1}$$

since B and C are independent. By question 2, B and C^c are independent. Therefore similarly we will get,

$$P(AC^c|B) = P(A|BC^c)P(C^c|B) = P(A|BC^c)P(C^c) \tag{2}$$

Hence,

$$\begin{aligned} P(A|B) &= P(AC|B) + P(AC^c|B) \\ &= P(A|BC)P(C) + P(A|BC^c)P(C^c) \\ &> P(A|B^cC)P(C) + P(A|B^cC^c)P(C^c). \end{aligned} \tag{3}$$

The last inequality holds because of the given conditions.

By 2.(a), B^c and C are independent. Therefore replacing B by B^C we will get that,

$$P(A|B^c) = P(AC|B^c) + P(AC^c|B^c) = P(A|B^cC)P(C) + P(A|B^cC^c)P(C^c) \tag{4}$$

By (3) and (4) we get,

$$P(A|B) > P(A|B^c).$$

- Suppose A and C are independent when conditioned on B. Then

$$P(AC|B) = P(A|B)P(C|B). \tag{5}$$

We have,

$$\begin{aligned}
& P(A|BC) > P(A|B^cC) \\
& \Rightarrow \frac{P(ABC)}{P(BC)} > \frac{P(AB^cC)}{P(B^cC)} \\
& \Rightarrow \frac{P(AC|B)P(B)}{P(C|B)P(B)} > \frac{P(AC|B^c)P(B^c)}{P(C|B^c)P(B^c)} \\
& \Rightarrow \frac{P(A|B)P(C|B)}{P(C|B)} > \frac{P(AC|B^c)}{P(C|B^c)}
\end{aligned}$$

Hence,

$$P(AC|B^c) < P(A|B)P(C|B^c). \tag{6}$$

Since $P(A|BC^c) > P(A|B^cC^c)$ and A and C^c are independent given B we can replace C by C^c in the above steps. we will similarly get

$$P(AC^c|B^c) < P(A|B)P(C^c|B^c). \tag{7}$$

Adding (6) and (7) we get,

$$P(A|B^c) < P(A|B).$$

- We now show that the conditions are not necessary. Let us go back to the scenario of part a. Let us assume that for country 1, the table is

	Smoker	Non-smoker
Death by lung cancer	21	10
No lung cancer	81	90

Therefore $P(A|BC) = .2059, P(A|B^cC) = .1$. Suppose the counts for country 2 remains the same. Then the combined table becomes

	Smoker	Non-smoker
Death by lung cancer	41	20
No lung cancer	161	180

i.e. $P(A|B) = .203, P(A|B^c) = .1$. Hence, $P(A|B) > P(A|B^c)$.
Now we have to check first whether B and C are independent. $P(BC) = P($ smoker and Country 1$) = \frac{102}{402} = \frac{17}{67} = .25373$. $P(B) = P(\text{Smoker}) = \frac{202}{402}$. $P(C) = P(\text{Country 1}) = \frac{202}{402}$. $P(B)P(C) = .25249 < P(BC)$. Hence, B and C are not independent.
Now, $P(AC|B) = \frac{21}{202} = .1039$. $P(A|B) = \frac{41}{202}$. $P(C|B) = \frac{102}{202}$. $P(A|B)P(C|B) = .1025 < P(AC|B)$. So A and C are not independent given B either.