Math 361 **Exam 1 Solutions and Comments** **Spring 2003**

Exam statistics:

(See the course web page for more details and grading information.)

Highest scores: 53 (2), 50 (3), 49, 48 (4)
Lowest scores: 22, 25 (2)
Average score: 40
Curve: Cutoffs for A/B/C/D were set at 45/38/30/20, respectively.

Problem 1.

Let A and B be independent events with probabilities $P(A) = 0.2$ and $P(B) = 0.3$. Let C denote the event "both A and B occur", and let D be the event "either A or B, but not both, occur."
(a) (4) Express D ("either A or B, but not both, occur") in terms of A and B using set-theoretic notation and compute $P(D)$.
(b) (3) Find $P(A|D)$.
(c) (3) Are C and D independent? Justify your answer. (3 points)

Solution.

(a) Since C and D are disjoint with positive probability (see below), they cannot be independent.
(a) $\boxed{D = (A \cup B) \setminus AB}$, or $\boxed{D = (A \setminus AB) \cup (B \setminus AB)}$. By the independence of A and B, we have $P(AB) = P(A)P(B)$, so $P(D) = P(A \cup B) - P(AB) = P(A) + P(B) - P(AB) - P(AB) = 0.2 + 0.3 - 2 \cdot 0.2 \cdot 0.3 = \boxed{0.38}$.
(b) $P(A|D) = P(AD)/P(D) = (P(A) - P(AB))/P(D) = (0.2 - 0.06)/0.38 = \boxed{7/19}$.
(c) The events C and D are disjoint, so $P(CD) = 0$, but $P(C)P(D) = 0.06 \cdot 0.38 \neq 0$, so C and D are $\boxed{\text{not independent}}$.
Remarks: By definition, independence of two sets C and D means that $P(CD) = P(C)P(D)$, and this is the condition that one has to check in order to determine whether two given sets are independent. In particular, independence is not the same as disjointness. In fact, in class it was shown that two disjoint events can only be independent if and one of them has probability 0.

Problem 2.

A box contains 100 numbered lottery tickets, of which 10 are winning tickets. You start drawing tickets, without replacement, one at a time, until you have found a winning ticket.
(a) (7) What is the probability that you need to draw exactly 5 tickets to obtain a winning ticket?
(b) (3) What is the probability that you need to draw exactly 5 tickets to obtain a winning ticket if the drawing is performed **with replacement**, (i.e., if you put each ticket taken out back into the box after you have checked whether it is a winning ticket)?

Solution.

(a) Label the tickets by $1 \dots, 100$ such that $1, \dots, 90$ denote the 90 non-winning tickets and $91, \dots, 100$ the 10 winning tickets. A suitable outcome space Ω is then given by the set of all tuples $(n_1, \dots, n_5)$ with $n_i = 1, \dots, 100$ and all n_i distinct (since the sampling is done without replacement), i.e.,

$$\Omega = \{(n_1, \dots, n_5) : n_i = 1, \dots, 100,\ n_i \text{ all distinct}\}.$$

Clearly, $\#(\Omega) = 100 \cdot 99 \cdots 96$. The event we are interested in, A="5th chip taken out is first good one", consists of those tuples in which $n_1, \dots, n_4$ are ≤ 90 (corresponding to the non-winning tickets) and distinct and n_5 is between 91 and 100 (the range corresponding to the winning tickets) i.e.,

$$A = \{(n_1, \dots, n_5) : n_1, n_2, n_3, n_4 = 1, \dots, 90; n_5 = 91, \dots, 100,\ n_i \text{ all distinct}\}.$$

Counting the number of elements of A by the slot method, we see that $\#(A) = 90 \cdot 89 \cdot 88 \cdot 87 \cdot 10$. Thus, the probability in question is

$$\boxed{P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{90 \cdot 89 \cdot 88 \cdot 87 \cdot 10}{100 \cdot 99 \cdots 96}}$$

(b) If the drawing is done with replacement, then the restriction to **distinct** n_i in the definition of A and Ω must be dropped, and the corresponding counts become $\#(A) = 90^4 \cdot 10$ and $\#(\Omega) = 100^5$. Thus,

$$\boxed{P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{90^4 \cdot 10}{100^5}}$$

[Alternatively, since sampling **with replacement** is equivalent to repeated S/F trials, the probability sought is that of the S/F sequence $FFFFS$, with $p = 10/100 = 0.1$, i.e., $(1-p)^4 p = 0.9^4 0.1$.]

Problem 3.

A student is taking a multiple choice exam in which each question has 5 possible answers, exactly one of which is correct. If the student knows the answer, she selects the correct answer. Otherwise she selects one answer at random from the 5 possible answers. Suppose that, for each question, there is a 70% chance that the student knows the answer.
(a) (5) State the general form of the average (or total probability) rule (involving events A and $B_1, \dots, B_n$) **along with any conditions that have to be satisfied for the rule to be applicable.** Then use this rule to compute the probability that, on a randomly chosen question, the student gets the correct answer.
(b) (5) State the general form of Bayes' rule (involving events A and $B_1, \dots, B_n$) **along with any conditions that have to be satisfied for the rule to be applicable**. Then use this rule to compute the probability that the student knows the answer to a question given that she has answered the question correctly.

Solution.

Relevant events: K="knows the answer"; A="gets the correct answer"
Given data: $P(K) = 0.7$ (since she knows the answer to 70 % of the questions); $P(A|K) = 1$ (if she knows the answer, she obviously gets the correct answer); $P(A|K^c) = 0.2$ (if she doesn't know the answer she gets the correct answer through guessing 1 out of 5 times, i.e., with probability 0.2).
To compute: The probability asked for in (a) is $P(A)$, that of part (b) is $P(K|A)$.
Computations:
(a) The average rule states that $P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$ **if the B_i's form a partition of Ω.** Applying this rule with A as above and sets $B_1 = K$ and $B_2 = K^c$ (which obviously form a partition of Ω) gives $P(A) = P(A|K)P(K) + P(A|K^c)P(K^c) = 1 \times 0.7 + 0.2 \times (1 - 0.7) = \boxed{0.76}$.
(b) Bayes' rule states that $P(B_i|A) = P(A|B_i)P(B_i)/P(A)$ (with $P(A)$ given by the average rule) **if the B_i's form a partition of Ω.**
Applying this rule with $B_1 = K$, and A as before, we get $P(K|A) = P(A|K)P(K)/P(A) = 1 \times 0.7/0.76 = \boxed{0.91}$.

Problem 4.

A lottery has a 1/300 chance of winning a prize.
(a) (5) If you play the lottery 200 times, what is the probability that you win at least twice? Give an exact answer. (Leave unsimplified.)
(b) (5) Give an approximation to the exact answer found in (a) (not involving any high powers or sums over a large number of terms).

Solution.

(a) Interpreting the 200 lottery games as success/failure trials with success probability $p = 1/300$, we need to compute the probability of getting at least two successes. Taking the complement, this becomes $1 - P(0) - P(1)$, where $P(0)$ and $P(1)$ are given by the binomial distribution with $n = 200$ and $p = 1/300$:

$$\boxed{1 - \binom{200}{0}\left(\frac{299}{300}\right)^{200} - \binom{200}{1}\left(\frac{1}{300}\right)\left(\frac{299}{300}\right)^{199}}$$

(b) Since $p = 1/300$ is small and of the order of $1/n = 1/200$, Poisson approximation is ideal for this situation. We have $\mu = np = 2/3$, so Poisson approximation gives for the above probability

$$\boxed{1 - e^{-2/3}\frac{(2/3)^0}{0!} - e^{-2/3}\frac{(2/3)^1}{1!} = 1 - \frac{5}{3}e^{-2/3}}$$

Problem 5.

A computer generates 10,000 random decimal digits $0, 1, 2, \ldots, 9$ (each selected with probability 1/10) and then tabulates the number of occurrences of each digit.

(a) (5) **Using an appropriate approximation**, find the (approximate) probability that exactly 1,060 of the 10,000 digits are 0. (Leave your answer in "raw" form such as $\pi^2 2^{-10}/200$.) (5 points)

(b) (5) Determine x (as small as possible) such that, with 99.95% probability, the number of digits 0 is at most x. (5 points)

Solution.

(a) We model the sequence of digits by a S/F trial sequence, with the trials corresponding to the 10,000 random digits and success meaning that the digit is 0. Thus $n = 10,000$ and $p = 1/10$. Normal approximation here is appropriate since n is very large and p only moderately small and since the probability sought falls near the center of the binomial distribution. We have $\mu = np = 10000$, $\sigma = \sqrt{10000 \cdot (1/10)(9/10)} = 30$, so normal approximation gives

$$P(1060 \text{ successes}) \approx \frac{1}{30}\phi\left(\frac{1060 - 1000}{30}\right) = \frac{1}{30}\phi(2) = \boxed{\frac{1}{30\sqrt{2\pi}}e^{-2}}.$$

(b) We need to choose x such that $P(\leq x \text{ successes}) = 0.99$. By normal approximation (with $\mu = 1000$ and $\sigma = 30$ from (a)), this probability is approximately equal to $\Phi((x - 1000)/30)$. Setting this equal to 0.9995, we find from the table, $(x - 1000)/30 = 3.3$, so $\boxed{x = 1099}$.

Problem 6 (Extracredit)

Find a good approximation to the binomial coefficient $\binom{2n}{n}$ for large n involving only elementary functions (such as $\binom{2n}{n} \approx 3n^2 2^{2n}$). (Hint: Look at $\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}$.) (5 points)

Solution.

The key here is to realize that the quantity given in the hint, i.e., $\binom{2n}{n}(1/2)^{2n}$, is the probability for getting n successes in $2n$ trials with success probability $p = 1/2$, and so can be approximated using normal approximation. With $\mu = (2n)(1/2) = n$ and $\sigma = \sqrt{(2n)(1/2)(1 - 1/2)} = \sqrt{n/2}$, we get

$$\binom{2n}{n}\left(\frac{1}{2}\right)^{2n} = P(n \text{ successes in } 2n \text{ trials}) \approx \frac{1}{\sigma}\phi\left(\frac{n - \mu}{\sigma}\right) = \frac{1}{\sqrt{n/2}}\phi(0) = \frac{1}{\sqrt{n\pi}},$$

so

$$\boxed{\binom{2n}{n} \approx \frac{2^{2n}}{\sqrt{n\pi}}}$$

Remark: The approximation given by this formula is surprisingly good, even when n is only moderately large. For example, for $n = 10$, $\binom{20}{10} = 184756$ and $2^{20}/\sqrt{10\pi} = 187079$, which is off by about 1%. For $n = 100$ the error is 0.1%.