

1.(a) The maximum-likelihood estimate of p is the observed relative frequency, i.e. $\hat{p} = X/N$.

(b) Since $E[X] = Np$ and $\text{var}(X) = Np(1-p)$, the Chebyshev inequality gives

$$P\{|\hat{p} - p| \geq 0.02\} = P\{|X - Np| > 0.02N\} \leq \frac{Np(1-p)}{0.0004N^2} = \frac{p(1-p)}{0.0004N} \leq \frac{1}{0.0016N}$$

where we have used the fact that the maximum value of $p(1-p)$ is 0.25 at $p = 0.5$. The right side of the above inequality must be no more than 0.05, and hence $N \geq 1/(0.0016 \cdot 0.05) = 12,500$. We conclude that polling 12,500 voters guarantees the desired confidence level regardless of the value of the parameter p .

2.(a) X is a binomial random variable with parameters (n, p) . $E[X] = np$. $\text{var}(X) = np(1-p)$.

(b) $P\{X > n/2\} = \sum_{i=(n+1)/2}^n \binom{n}{i} p^i (1-p)^{n-i}$. For $n = 3$, $p = 10^{-2}$, this is $3p^2(1-p) + p^3 = 3p^2 - 2p^3 \approx 3 \times 10^{-4}$

Repetition *does* reduce the probability of error, but also reduces the data rate by a factor of 3 (n generally).

(c) $P\{X > n/2\} = P\{(X-np) > n/2 - np\} \leq P\{|X-np| > n/2 - np\} \leq \frac{np(1-p)}{[n(1/2-p)]^2} = \frac{p(1-p)}{n(0.5-p)^2} = \frac{4.1232 \times 10^{-2}}{n}$
which is less than 10^{-5} for $n > 4123$.

3.(a) $P\{X \leq 100\} = 1 - P\{X > 100\} = 1 - [P\{X = 105\} + P\{X = 104\} + P\{X = 103\} + P\{X = 102\} + P\{X = 101\}]$
 $= 1 - \left[(0.9)^{105} + \binom{105}{1}(0.9)^{104}(0.1) + \binom{105}{2}(0.9)^{103}(0.1)^2 + \binom{105}{3}(0.9)^{102}(0.1)^3 + \binom{105}{4}(0.9)^{101}(0.1)^4 \right]$
 $= 0.983283...$

(b) If X is a binomial random variable with parameters (n, p) , then $Y = n - X$ is a binomial random variable with parameters $(n, 1-p)$. Thus, the number of no-shows is a binomial random variable with parameters $(105, 0.1)$. Since n is large and p is small, it is reasonable to approximate this as a Poisson random variable with parameter $\lambda = np = 10.5$.

(c) $\{Y \geq 5\} = 1 - P\{Y < 5\} = 1 - [P\{Y = 0\} + P\{Y = 1\} + P\{Y = 2\} + P\{Y = 3\} + P\{Y = 4\}]$
 $= 1 - \exp(-10.5)[1 + (10.5) + (10.5)^2/2 + (10.5)^3/6 + (10.5)^4/24] = 0.978906...$ which is not bad....

(d) Let Y denote the number of passengers who show up for the flight, regardless of where they came from. When the connecting flight is on time, 15 passengers are guaranteed to show up. Of the other 90, the number who show up is a binomial random variable Z with parameters $(90, 0.9)$. Thus, $Y = 15 + Z$ when the connecting flight is on time. Otherwise, when the connecting flight is late, only Z passengers show up, and hence $Y = Z$ in this case.

Thus, $P\{Y \leq 100 \mid \text{on time}\} = P\{Z \leq 85\} = 1 - P\{Z > 85\} = 0.95345...$ and $P\{Y \leq 100 \mid \text{late}\} = 1$.

Hence, $P(Y \leq 100) = (0.95345...) \times (1/3) + 1 \times (2/3) = 0.98448...$. Note that the only way for the airline to better the chances of everyone getting a seat is to increase the probability that the connecting flight is late!

4.(a) Two boxes, of course!

(b) For $n \geq 2$, $P\{X = n\} = P\{\text{LLL...LD}\} + P\{\text{DDD...DL}\} = (2/3)^{n-1}(1/3) + (1/3)^{n-1}(2/3) = (2^{n-1}+2)/3^n$.

$$\sum_{n=2}^{\infty} (2^{n-1}+2)/3^n = \frac{2}{9} \cdot \frac{e}{e-1} + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \dots + \frac{2}{9} \cdot \frac{e}{e-1} + \frac{1}{3} + \left(\frac{1}{3}\right)^2 + \dots = \frac{2}{9} \left[\frac{1}{1-2/3} + \frac{1}{1-1/3} \right] = 1.$$

(c) If the first box contains a picture of Luke, then Jimmy has to wait for a box with Darth. Thus, the number of *additional* boxes is a geometric random variable Y with parameter $1/3$. Conversely, if the first box contains a picture of Darth, then Jimmy has to wait for a box with Luke. Thus, the number of *additional* boxes is a geometric random variable Z with parameter $2/3$. Thus, for $n \geq 2$,

$$P\{X = n \mid \text{first box is L}\} = P\{Y = n-1\} = (2/3)^{n-2}(1/3) \quad (\text{note that } X = 1 + Y)$$

$$\text{and } P\{X = n \mid \text{first box is D}\} = P\{Z = n-1\} = (1/3)^{n-2}(2/3). \quad (\text{note that } X = 1 + Z)$$

We note for later use that the conditional mean and variance of X are given by

$$E[X \mid L] = E[1+Y] = 1 + E[Y] = 1 + 1/(1/3) = 4; \text{ var}(X \mid L) = \text{var}(1+Y) = \text{var}(Y) = (1 - 1/3)/(1/3)^2 = 6;$$

$$E[X \mid D] = E[1+Z] = 1 + E[Z] = 1 + 1/(2/3) = 5/2; \text{ var}(X \mid D) = \text{var}(1+Z) = \text{var}(Z) = (1 - 2/3)/(2/3)^2 = 3/4.$$

(d) $P\{X = n\} = P\{X = n \mid \text{first box is L}\}P\{\text{first box is L}\} + P\{X = n \mid \text{first box is D}\}P\{\text{first box is D}\}$
 $= (2/3)^{n-2}(1/3)(2/3) + (1/3)^{n-2}(2/3)(1/3) = (2^{n-1}+2)/3^n$ just as in part (b).

(e) $E[X] = E[X \mid \text{first box is L}] \cdot P\{\text{first box is L}\} + E[X \mid \text{first box is D}] \cdot P\{\text{first box is D}\}$
 $= 4 \cdot (2/3) + (5/2) \cdot (1/3) = 7/2$. Note that the unconditional mean is between the two conditional means...
 $\text{var}(X)$ is the sum of the mean of the conditional variance $(6 \cdot (2/3) + (3/4) \cdot (1/3))$ and the variance of the

conditional mean which is readily calculated as $(4^2 \cdot (2/3) + (5/2)^2 \cdot (1/3)) - (7/2)^2$ (huh? what formula is that?) or as $(4 - 7/2)^2 \cdot (2/3) + (5/2 - 7/2)^2 \cdot (1/3)$ giving us that $\text{var}(\mathbf{X}) = 5.75$

5.(a) $P(A|B^C) = 1 - P(A \cap B^C) = 0.6$. $P(A) = P(A|B)P(B) + P(A|B^C)P(B^C) = 0.39$.
 $P(B|A) = P(AB)/P(A) = P(A|B)P(B)/P(A) = 0.21/0.39 = 7/13$.

(b) $P(F) = P(EF)/P(E|F) = P(F|E)P(E)/P(E|F) = (1/2) - (1/4)/(1/3) = 3/8$.

(c) $1 \geq P(G \cup H) = P(G) + P(H) - P(GH)$.

Hence, $P(GH) \geq 2/3 + 2/3 - 1 = 1/3$, and $P(G|H) = P(GH)/P(H) = P(GH)/(2/3) \geq (1/3)/(2/3) = 1/2$.

6.(a) $P(R_1) = 6/10$, $P(R_2|R_1) = 5/9$, $P(R_2|R_1^C) = 6/9$. Geez, that was an easy one!

(b) $P(R_2) = P(R_2|R_1)P(R_1) + P(R_2|R_1^C)P(R_1^C) = (5/9) \cdot (6/10) + (6/9) \cdot (4/10) = 54/90 = 6/10$. Surprised?

(c) $P(R_1) = 6/10$ as before; $P(R_2|R_1) = 9/13$, $P(R_2|R_1^C) = 6/13$.

$P(R_2) = P(R_2|R_1)P(R_1) + P(R_2|R_1^C)P(R_1^C) = (9/13) \cdot (6/10) + (6/13) \cdot (4/10) = 78/130 = 6/10$. Surprised **now**?

A ball is being picked at random from an urn with 13 balls in it. It would seem that **all** probabilities would be of the form $k/13$, but instead we get an answer of $6/10$. This is called Polya's urn scheme. Note that the answer is $6/10$ if *any* number of additional balls of the same color are added; the result is not a magical property of 3. Furthermore, if the second ball is returned to the urn with 3 additional balls of the same color, the probability of the third ball being drawn being red is *still* $6/10$. Weird, isn't it?

Noncredit Exercises:

1. $(d/d\lambda)P\{X = k\} = (d/d\lambda)\exp(-\lambda)\lambda^k/k! = \exp(-\lambda)[k\lambda^{k-1}/k! - \lambda^k/k!] = 0$ at $\lambda = k$. Since $\exp(-\lambda)\lambda^k/k! > 0$ for all $\lambda \in (0, \infty)$ and has value 0 at $\lambda = 0$ and also at $\lambda = \infty$, it must have a maximum at $\lambda = k$. Thus the maximum-likelihood estimate of λ is $\hat{\lambda} = k$. Since $\lambda \approx Np$, this is consistent with the binomial estimate $\hat{p} = k/N$.

2.(a) $\sum_{k \text{ even}} P\{X = k\} = \exp(-\lambda) \sum_{k=0}^{\infty} \lambda^{2k}/2k! = \exp(-\lambda) \cosh(\lambda)$ via the known series for cosh. Geez, that was easy!

(b) $[1 + (1 - 2p)^N]/2 = [1 + (1 - 2np/n)^N]/2 \approx [1 + \exp(-2np)]/2 = \exp(-np)[\exp(np) + \exp(-np)]/2$
 $= \exp(-np) \cosh(np) = \exp(-\lambda) \cosh(\lambda)$ on setting $\lambda = np$. He's losing his touch; that was even easier!

3. Let A and B denote respectively the events that your **first** choice and your **final** choice is the curtain concealing the prize. $P(A) = 1/3$, $P(A^C) = 2/3$.

(a) If you always switch, $P(B|A) = 0$, while $P(B|A^C) = 1$. Hence, $P(B) = P(B|A)P(A) + P(B|A^C)P(A^C) = 2/3$.

(b) If you never switch, then $P(B|A) = 1$, while $P(B|A^C) = 0$. Hence, $P(B) = P(B|A)P(A) + P(B|A^C)P(A^C) = 1/3$.

(c) If you decide at random, then $P(B|A) = P(B|A^C) = 1/2$ and $P(B) = P(B|A)P(A) + P(B|A^C)P(A^C) = 1/2$ also. Monty is correct in his assertion. (Would he lie to you? Besides, you saw it on **TV**, so it must be true!!!!)

4. This game is different from the one in Problem 3 in that *you have no idea what the rules of the game are*. If the man is intent on separating you from your money as quickly as possible, he will not offer the chance to switch unless you picked the shell hiding the pea in the first place! That is, if you picked the wrong shell, the man will reveal the pea and you will lose your bet. Of course, if you look like a person willing to play several rounds, the man may set you up by playing by Monty's rules (and allowing you to win with probability $2/3$) for some time. Then you will place a large bet on the wrong shell, and all of a sudden, you will not be given the choice of changing your bet! Personally, I would stick with the shell originally chosen since it gives me at least a $1/3$ probability of winning regardless of the man's strategy; your experience (and monetary losses) may vary