

- 1.(a) $\mathbf{Y}$ is a binomial random variable with parameters $(N, 1/2)$. $p_Y(k) = P\{\mathbf{Y} = k\} = \binom{N}{k} \left(\frac{1}{2}\right)^N$ for $0 \leq k \leq N$,

and we can write without further computation that $E[\mathbf{Y}] = N \cdot (1/2) = N/2$.

- (b) Each week, the investment value is multiplied by either $3/2$ or $1/2$. Hence, after N weeks, an investment of $\$C$ has value $\$C \cdot (3/2)^Y (1/2)^{N-Y} = \$C \cdot 3^{Y/2} / 2^{N/2} = \$C \cdot (3/4)^{N/2} < \$C$ if $\mathbf{Y}$ has value $N/2$, that is, half the weeks gave a 50% gain and half a 50% loss. So, equal numbers of gains and losses results in a net loss. MBAs would, of course, say that the average gain of the GoGoDotCom investments over N weeks was $(50\% + 50\% - 50\% + 50\% \dots) / N = 0\%$!! since $N/2$ weeks has gains and $N/2$ weeks losses.

- (c) $1 + \mathbf{Z} = (1 + \mathbf{X}_1)(1 + \mathbf{X}_2)(1 + \mathbf{X}_3) \dots (1 + \mathbf{X}_N) = (3/2)^Y (1/2)^{N-Y} = 3^Y / 2^N$. Hence, $\mathbf{Z} = 3^Y / 2^N - 1$ takes on values $2^{-N} - 1, 3 \cdot 2^{-N} - 1, 3^2 \cdot 2^{-N} - 1, 3^3 \cdot 2^{-N} - 1, \dots, 3^N \cdot 2^{-N} - 1$. Note that $\mathbf{Z} < 0$ for $\mathbf{Y} < N/\log_2 3$.

- (d) The pmf of $\mathbf{Z}$ is given by $P\{\mathbf{Z} = 3^k \cdot 2^{-N} - 1\} = \binom{N}{k} \left(\frac{1}{2}\right)^N$ for $0 \leq k \leq N$. From this, it follows that

$$E[\mathbf{Z}] = \sum_{k=0}^N (3^k \cdot 2^{-N} - 1) \cdot \binom{N}{k} \left(\frac{1}{2}\right)^N = 2^{-N} \cdot (1 + 3)^N - 1 = 0 \text{ !!}$$

Thus, the average gain is in fact 0, but note that 0 is not one of the values taken on by $\mathbf{Z}$; that is, *no* combination of gains and losses will leave you with exactly the same amount that you began with: you will either have made money or lost money by investing with GoGoDotCom.

- (e) $P\{\mathbf{Z} < 0\} = P\{\mathbf{Y} < N/\log_2 3\}$ must be computed numerically; there is no closed-form expression for it. Since $N/\log_2 3 > N/2$, it is easier to compute the probability $P\{\mathbf{Y} > N/\log_2 3\} = P\{\mathbf{Y} > 8.2\} = P\{\mathbf{Y} \geq 9\} = P\{\mathbf{Y} \leq 4\}$. Why? We have by calculation (or by consulting, for example, *Standard Mathematical Tables and Formulae*, CRC Press, 1996 (p. 653)) that $P\{\mathbf{Y} \leq 4\} = 0.1334$. Thus, you have probability 0.8666 of losing money on this investment even though the average gain is 0 !!

2. $P\{\mathbf{X} \text{ is even}\} = P\{\mathbf{X} = 0\} + P\{\mathbf{X} = 2\} + \dots + P\{\mathbf{X} = M\}$ where $M = N$ if N is even and $M = N - 1$ if N is odd. Hence, $P\{\mathbf{X} \text{ is even}\} = q^N + \binom{N}{2} q^{N-2} p^2 + \dots + \binom{N}{M} q^{N-M} p^M$. Note that the last term is p^N if N

is even, and it is Nqp^{N-1} if N is odd (whatever happened to $\binom{N}{M}$??)

Now, $(q + p)^N = q^N + \binom{N}{1} q^{N-1} p + \binom{N}{2} q^{N-2} p^2 + \dots + \binom{N}{N-1} qp^{N-1} + p^N$ while

$$(q - p)^N = q^N - \binom{N}{1} q^{N-1} p + \binom{N}{2} q^{N-2} p^2 - \dots + (-1)^{N-1} \binom{N}{N-1} qp^{N-1} + (-1)^N p^N.$$

We add these equations together and note alternate terms cancel. Also if N is odd, the last terms cancel, while if N is even, the next-to-last terms cancel. In either case, we see that the sum is just $2P\{\mathbf{X} \text{ is even}\}$, i.e. $P\{\mathbf{X} \text{ is even}\} = [(q+p)^N + (q-p)^N] / 2 = [1 + (1 - 2p)^N] / 2$. Exercise: when is $P\{\mathbf{X} \text{ is even}\} < 1/2$?

- 3.(a) $P\{\mathbf{W} = n\} = \binom{N-K}{n} (0.8)^n (0.2)^{N-K-n}$, for $n = 0, 1, 2, \dots, N-K$.

- (b) For fixed N and n , the answer to part (a) is a function of K , say $f(K)$. Then,

$$\frac{f(K)}{f(K-1)} = \frac{\binom{N-K}{n} (0.8)^n (0.2)^{N-K-n}}{\binom{N-(K-1)}{n} (0.8)^n (0.2)^{N-(K-1)-n}} = \frac{(N-K)(N-K-1) \dots (N-K-n+1)}{(N-(K-1))(N-(K-1)-1) \dots (N-(K-1)-n+1) 0.2} \geq 1 \text{ if and only}$$

if $\frac{N-K-n+1}{N-K+1} \geq 0.2$, that is, $K \leq N - 1.25n + 1$. Since K is an integer, the maximum-likelihood estimate

can be expressed as $\hat{K} = \lfloor N - 1.25n + 1 \rfloor$. For $N = 100$ and $n = 8$, $N - 1.25n + 1 = 91$ is an integer.

Hence, $f(91)/f(90) = 1$ and either 90 or 91 can be taken to be the maximum-likelihood estimate of K . The formula we derived gives 91 as the estimate but it is perfectly legitimate to use 90 instead.

- (c) The examiner's estimate $\hat{K}$ is $C - \lceil 0.25n \rceil = N - n - \lceil 0.25n \rceil$ which has value 90 if $N = 100$ and $n = 8$ while $\hat{K}$ is 91. Thus, the examiner's estimate is smaller than the answer obtained in part (b), but still perfectly defensible as a maximum-likelihood estimate as described in part (b) (she is just a tough grader!). In fact, $\hat{K} = \hat{K}$ except when $0.25n$ happens to be an integer (because $N - n - \lceil 0.25n \rceil = \lfloor N - 1.25n + 1 \rfloor$ except when $0.25n$ is an integer) and in this special case, the examiner is opting for the lower value.

For the case $N = 100$ and $n = 10$, we get $\hat{K} = \lfloor 100 - 12.5 + 1 \rfloor = \lfloor 88.5 \rfloor = 88$ while $\hat{K}$ is $C - \lceil 0.25n \rceil = 90 - \lceil 2.5 \rceil = 90 - 3 = 87$ is the same. The unlucky student who blew all 10 guesses suffers the further indignity of having a score of 88 on the exam even though the answers to 90 questions were known (and recorded correctly on the answer sheet!)

- (d) If $N = 100$ and $K = 90$, $\mathbf{W}$ is a binomial random variable with parameters $(10, 0.8)$ and hence $P\{\mathbf{W} = n\}$ is maximized at $n = \lceil 0.8 \times (10+1) \rceil = 8$. If this maximum-probability event actually occurs, then as discussed in parts (b) and (c), $\hat{K} = 90$ which correctly estimates K , while $\hat{K} = 91$ slightly overestimates K .
- (e) If $\mathbf{W} = 4$, the guessing penalty reduces the score by just 1 — the student gets a 95 while knowing only 90 answers whereas the unlucky one who blew all 10 guesses got only an 88 while knowing 90 answers! This is the origin of the saying “Some people have all the luck!”
- (f) It is easily verified that if C equals 100, 99, 98, 97, 96, 95, 94, 93, 92, 91, and 90, then the examiner’s estimate $\hat{K}$ is 100, 99, 98, 97, 95, 94, 93, 92, 90, 89, and 88. The probability of underestimating K is $P\{\mathbf{W} = 9\} + P\{\mathbf{W} = 10\} = 0.37580964\dots$ while the probability of correctly estimating K is $P\{\mathbf{W} = 8\} = 0.30198989\dots$ The probability of overestimating K can be easily found as $1 - P\{\text{correctly estimating}\} - P\{\text{underestimating}\} = 0.32220047\dots$ It can also be computed directly (and with a lot more work!) as $P\{\mathbf{W} \leq 7\} = P\{\mathbf{W} = 0\} + P\{\mathbf{W} = 1\} + \dots + P\{\mathbf{W} = 7\} = 0.32220047\dots$ More interestingly, write the examiner’s estimate as $\hat{K} = N - \mathbf{W} - \lceil 0.25\mathbf{W} \rceil$ and think of it as a random variable. The average value of $\hat{K}$, i.e. $E[\hat{K}]$, is $90.3230648\dots$ which is pretty good estimation! It is interesting to note that $N - \mathbf{W} - 0.25\mathbf{W}$ (without the floor function) has average value precisely K , so the examiner is contributing to grade inflation while being a tough grader at the same time!

4.(a) $\mathbf{X}$ takes on values $-6, 6, 12, 18$.

(b) If \$6 is bet on i , then you lose it if all three dice show one of the 5 non- i numbers.

Hence $P\{\mathbf{X} = -6\} = 5^3/6^3 = \frac{125}{216}$. On the other hand, you win \$6 if one of the three dice shows i and the

other two have non- i numbers. Hence, $P\{\mathbf{X} = 6\} = 3 \cdot (1 \cdot 5^2)/6^3 = \frac{75}{216}$. By a similar argument, $P\{\mathbf{X} = 12\}$

is $3 \cdot (1^2 \cdot 5)/6^3 = \frac{15}{216}$, and $P\{\mathbf{X} = 18\} = 1^3/6^3 = \frac{1}{216}$. Sanity check: $125 + 75 + 15 + 1 = 216$, so we have not left anything out.

(c) $E[\mathbf{X}] = \sum u \cdot p(u) = \frac{125 \cdot (-6) + 75 \cdot 6 + 15 \cdot 12 + 1 \cdot 18}{216} = -\frac{102}{216} = -\frac{17}{36}$ i.e. a loss of roughly 47¢ per game.

(d) The chance of the three dice showing three different numbers is $\frac{6 \cdot 5 \cdot 4}{216} = \frac{5}{9}$. In this case, you come out even since you win \$3 on the three numbers showing, but lose \$3 on the three no-shows. The chance that the three dice show the same number is $6 \cdot \frac{1}{216} = \frac{1}{36}$ in which case you win \$3 on the winning number but lose \$5 on the 5 no-shows for a net loss of \$2. The probability that exactly two numbers are identical is thus $\frac{15}{36} = \frac{5}{12}$ in which case, you win \$2 on the pair and \$1 on the singleton, but lose \$4 on the other no-shows for a net loss of -1 . Thus, $\mathbf{Y}$ is a random variable taking on values $0, -1, -2$, with probabilities as found above. (Remind me once again why you are bothering to play this game at all?), and its expected value is $E[\mathbf{Y}] = \frac{0 \cdot 20 - 1 \cdot 15 - 2 \cdot 1}{36} = -\frac{17}{36}$ just as before. Splitting your bet six ways has no effect on your losses!

Exercise: Would it be better to bet on 2 or 3 or 4 or 5 numbers (in equal shares) instead?

5. There are $\binom{10}{5}$ different ways of rank-ordering 5 men and 5 women, and all these are equally likely as stated in the problem. The random variable $\mathbf{X}$, the ranking of the highest-ranked woman, obviously can take on values 1, 2, 3, 4, 5, 6 only. Now, if $\mathbf{X} = k$, the other 4 women can be in placed in the lower $10-k$ ranks in $\binom{10-k}{4}$ different ways. (Sanity check: the other $5 - (k-1) = 6-k$ men can be placed in the lower $10-k$ ranks in $\binom{10-k}{6-k} = \binom{10-k}{4}$ different ways where we have applied the result: $\binom{n}{i} = \binom{n}{n-i}$.)

Hence, for $1 \leq k \leq 6$, $P\{\mathbf{X} = k\} = \frac{\binom{10-k}{4}}{\binom{10}{5}} = 5 \cdot \frac{(10-k) \cdot (10-k-1) \cdot (10-k-2) \cdot (10-k-3)}{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}$. The numerical values

are respectively $\frac{1}{2}, \frac{5}{18}, \frac{5}{36}, \frac{5}{84}, \frac{5}{252}$, and $\frac{1}{252}$ for $k = 1, 2, 3, 4, 5, 6$. Sanity check: The numbers add to 1.

6.(a) If $\mathbf{X} \geq 50$, he sells all the papers. This event occurs with probability $P\{\mathbf{X} \geq 50\} \approx 0.983$

- (b) If he cannot sell any papers, he loses $\$0.1 \times 50 = \$5$. If he sells all of them, he makes $\$0.25 \times 50 = \$12.50$. If the demand X is for 50 papers or less, the newsboy sells X papers and recycles $50 - X$ papers. Thus, his profit (in cents) is $25X - 10(50 - X) = 35X - 500$. Otherwise, $X > 50$, and his profit is 1250 regardless of how large X is. Hence, Z is related to X as $Z = \begin{cases} 35X - 500, & 0 \leq X \leq 50, \\ 1250, & X > 50. \end{cases}$
- (c) Z takes on values $-500, -465, -430, \dots, 1215, 1250, 1250, \dots, 1250$ as X takes on values $0, 1, 2, \dots, 49, 50, 51, \dots, 100$. Thus, $P\{Z = -500\} = P\{X = 0\}$, etc. and we can thus evaluate $E[Z]$ via LOTUS as $E[Z] = (-500)P\{X = 0\} + (-465)P\{X = 1\} + \dots + 1215P\{X = 49\} + 1250P\{X = 50\} + 1250P\{X = 51\} + \dots + 1250P\{X = 100\}$. The last 51 terms can be combined to write $E[Z] = (-500)P\{X = 0\} + (-465)P\{X = 1\} + \dots + 1215P\{X = 49\} + 1250P\{X \geq 50\} \approx 1248.63$. This is not too surprising since there is a better than 98% chance he sells all his papers and makes 1250 every day. (It *would have* been surprising if the average profit had been *more* than 1250, the max profit!)
- (d) The 51st paper provides an *extra* profit of 25 cents if $X > 50$ (which occurs with probability ≈ 0.9729) and a loss of 10 cents if $X \leq 50$ (which occurs with probability ≈ 0.0271). Hence, the average **additional** profit from the 51st paper is $A(50) \approx 25 \times 0.9729 - 10 \times 0.0271 \approx 24.05$ and the total average profit from 51 papers is approximately $1248.63 + 24.05 = 1272.68$.
- (e) The 52nd paper provides an *extra* profit of 25 cents if $X > 51$ (which occurs with probability ≈ 0.958) and a loss of 10 cents if $X \leq 51$ (which occurs with probability ≈ 0.042). Hence, the average **additional** profit from the 52nd paper is $A(51) \approx 25 \times 0.958 - 10 \times 0.042 \approx 23.52$ and the total average profit from 52 papers is approximately $1272.68 + 23.52 = 1296.20$. Notice that on average he makes **less additional** money (23.52 cents) from the 52nd paper than he does from the 51st (24.05 cents).
- (f) Generally, the $(N+1)$ th paper is sold for 25 cents profit if $X > N$, and remains unsold for 10 cents loss if $X \leq N$. Hence, $A(N) = 25 \times P\{X > N\} - 10 \times P\{X \leq N\}$. But, $P\{X > N\} + P\{X \leq N\} = 1$, so we can write that $A(N) = 25 - 35 \times P\{X \leq N\}$. But, $P\{X \leq N\} = P\{X = 0\} + P\{X = 1\} + \dots + P\{X = N\}$ increases as N increases (since we keep adding in more and more positive terms as N increases) to value 1 when $N = 100$. Hence, $A(N)$ **decreases** as N increases, and it becomes negative when N gets to be so large that $P\{X \leq N\}$ exceeds $25/35 = 0.71\dots$. Note that this **must** eventually happen for some N since $P\{X \leq 100\} = 1$.
- (g) $A(50) \approx 24.05 > 0$, so that buying 51 papers gives more average profit than buying 50 papers, and buying 52 papers gives even more profit since $A(51) > 0$. Since $P\{X \leq 62\} < 25/35$, $A(62) > 0$, and thus buying 63 papers increases profits even further. However, $P\{X \leq 63\} > 25/35$ and hence $A(63) < 0$, i.e. the **64th** paper has a loss on average: the newsboy should buy 63 papers to sell (just a little bit more than the average daily demand of 60 papers. His average profit is approximately 1442.06. (With 64 papers, the profit is approximately 1440.41 cents which is a tad smaller, and profits decrease further as more papers are bought.)

Exercise: what should the newsboy do if the numbers had turned out such that $P\{X \leq N\} = 25/35$ *exactly* for some N , so that $A(N) = 0$? Should he buy N papers or $N+1$?