Kevin Chau
23816929
CS70-2H

Homework 14

1. Alice trying to send Bob m packets
probability p of a packet error
Alice sends $n > m$ packets
Bob to decode with probability r

a) Bob cannot decode the message if
$m+2k > n$, where $k = \#$ of corrupted packets

$1-r = P(\text{cannot decode})$
$= P(m+2k > n) = P(2k > n-m) = \boxed{\sum_{i=\frac{n-m}{2}}^{n} \binom{n}{i} p^i (1-p)^{n-i}}$
$= \sum_{i=\frac{n-m}{2}}^{n-m} P(k=i)$

$k \equiv \sum_{i=0}^{n} x_i$ where y_i is 1 if the ith packet is corrupt

since the number of corrupt packets in an undecodable message is between $\frac{n-m}{2}$ and n errors. The probability of getting i errors from modeling the errors as a biased binomial coin:

$$P(k=i) = \binom{n}{i} p^i (1-p)^{n-i}$$

b) $r = 1 - P(\text{cannot decode}) = 1 - \sum_{i=\frac{n-m}{2}}^{n} \binom{n}{i} p^i (1-p)^{n-i} = .9$

$p = .1 \quad n = 100$

$P(m+2k > n) \leq \frac{\text{var}(2k)}{n} = \frac{4\,\text{var}(k)}{n}$

$\sum_{i=50-\frac{m}{2}}^{100} \binom{100}{i} (.1)^i (.9)^{100-i} \leq \frac{4np(1-p)}{n^2} = \frac{4(.1)(.9)}{100}$

2. X is the random variable representing total number of heads in n coin tosses.

a) let x_i be 0 if the ith toss is tails, and 1 if the ith toss is heads. then

$$X = \sum_{i=1}^{n} x_i$$

$$\mu = E[X] = E\left[\sum_{i=1}^{n} x_i\right]$$

using linearity of expectations:

$$\mu = \sum_{i=1}^{n} E[x_i] = n \cdot E[x_1]$$

$$E[x_1] = \tfrac{1}{2}(0) + \tfrac{1}{2}(1) = \tfrac{1}{2}$$

$$\Rightarrow \boxed{\mu = n \cdot \tfrac{1}{2} = \tfrac{n}{2}} = E[X]$$

$$\mathrm{var}[X] = E[X^2] - E[X]^2 = E\left[\left(\sum_{i=1}^{n} x_i\right)^2\right] - \frac{n^2}{4}$$

$$= E\left[\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n}\sum_{(i \neq j)} x_i x_j\right] - \frac{n^2}{4}$$

$$= n E[x_1^2] + (n^2 - n) E[x_1 x_2] - \frac{n^2}{4}$$

$$= n\left(\tfrac{1}{2}(0) + \tfrac{1}{2}(1)\right) + (n^2 - n)\left(\tfrac{3}{4}(0) + \tfrac{1}{4}(1)\right) - \frac{n^2}{4}$$

$$= \frac{n}{2} + \frac{n^2 - n}{4} - \frac{n^2}{4} = \frac{2n + \cancel{n^2} - n - \cancel{n^2}}{4} = \boxed{\frac{n}{4} = \mathrm{var}[X]}$$

b) Prove $P(X = k) = \binom{n}{k} / 2^n$

For n coin tosses, there are 2^n different binary strings that each represent a possible outcome (ie 10100011...)

Since every outcome has equal probability of $(\frac{1}{2})^n$, the distribution is uniform, so the probability is just a counting problem. There are $\binom{n}{k}$ ways for our binary outcome string to have k heads (a length n binary string - choose k places for 1s)

Then $\boxed{P(X = k) = \binom{n}{k} / 2^n}$

c) Stirling's formula: $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$

$$P(X=k) = \frac{\binom{n}{k}}{2^n} = \frac{n!}{(n-k)!\,k!\,2^n} \approx \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi(n-k)}\left(\frac{n-k}{e}\right)^{n-k}\sqrt{2\pi k}\left(\frac{k}{e}\right)^k 2^n}$$

$$\approx \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(n-k)k}}\;\frac{\cancel{\left(\frac{1}{e}\right)^n} n^n}{(n-k)^n(n-k)^{-k}\cancel{\left(\frac{1}{e}\right)^n}\cancel{\left(\frac{1}{e}\right)^{-k}} k^k \cancel{\left(\frac{1}{e}\right)^k} 2^n}\cdot\frac{2^{-k}}{2^{-k}}\cdot\frac{n^{-k}}{n^{-k}}$$

$$\approx \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(n-k)k}}\;\frac{n^{n-k}}{2^{n-k}(n-k)^{n-k}}\;\frac{n^k}{2^k k^k}$$

$$\boxed{P(X=k) \approx \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{(n-k)k}}\left(\frac{n}{2(n-k)}\right)^{n-k}\left(\frac{n}{2k}\right)^k}$$

$$\Rightarrow P(X=k) \approx \frac{1}{\sqrt{2\pi}}\left(\frac{n}{2(n-k)}\right)^{n-k}\left(\frac{n}{2k}\right)^k \frac{2}{\sqrt{n}}$$

$$\frac{2n-2-n}{\sqrt{n}}$$

d) $Y = (X-\mu)/\sigma = \frac{X-\frac{n}{2}}{\sqrt{n/4}} = \frac{2X-n}{\sqrt{n}}$ $\qquad \frac{2}{\sqrt{n}} - \frac{n}{\sqrt{n}}$

$Y_{max} = \frac{2n-n}{\sqrt{n}} = \frac{n}{\sqrt{n}}$ $\qquad Y_{min} = -\frac{n}{\sqrt{n}}$ $\qquad Y = \left\{-\frac{n}{\sqrt{n}}, \frac{2}{\sqrt{n}}-\frac{n}{\sqrt{n}}, \frac{4}{\sqrt{n}}-\frac{n}{\sqrt{n}}, \ldots\ldots, \frac{n}{\sqrt{n}}-\frac{2}{\sqrt{n}}, \frac{n}{\sqrt{n}}\right\}$

$\boxed{d = \frac{2}{\sqrt{n}}}$ $\qquad \Delta Y = \frac{\cancel{2X}+2-n}{\sqrt{n}} - \frac{\cancel{2X}-n}{\sqrt{n}} = \frac{2}{\sqrt{n}}$

e) $\frac{P(Y=t)}{d} = \frac{P(X=t)}{d}$

$P(X=k) = \binom{n}{k}/2^n$ $\qquad t = Y(k) = \frac{k-\mu}{\sigma} = \frac{k-\frac{n}{2}}{\sqrt{n}/2} = \frac{2k-n}{\sqrt{n}}$

$2k = \sqrt{n}\,t + n \Rightarrow k = \frac{n}{2}\left(\frac{t}{\sqrt{n}}+1\right)$

$$\frac{P(Y=t)}{d} \approx \frac{1}{\sqrt{2\pi}}\,\frac{\cancel{\sqrt{n}}}{2}\cdot\frac{\cancel{2}}{\cancel{\sqrt{n}}}\left(\frac{n}{2\left(n-\frac{n}{2}\left(\frac{t}{\sqrt{n}}+1\right)\right)}\right)^{n-\frac{n}{2}\left(\frac{t}{\sqrt{n}}+1\right)}\left(\frac{n}{2\frac{n}{2}\left(\frac{t}{\sqrt{n}}+1\right)}\right)^{\frac{n}{2}\left(\frac{t}{\sqrt{n}}+1\right)}$$

$$= \frac{1}{\sqrt{2\pi}}\left(\frac{n}{2n-n\left(\frac{t}{\sqrt{n}}+1\right)}\right)^{\frac{n}{2}\left(1-\frac{t}{\sqrt{n}}\right)}\left(\frac{1}{\left(\frac{t}{\sqrt{n}}+1\right)}\right)^{\frac{n}{2}\left(\frac{t}{\sqrt{n}}+1\right)}$$

$$= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{1-\frac{t}{\sqrt{n}}}\right)^{\frac{n}{2}\left(1-\frac{t}{\sqrt{n}}\right)}\left(\frac{1}{1+\frac{t}{\sqrt{n}}}\right)^{\frac{n}{2}\left(\frac{t}{\sqrt{n}}+1\right)}$$

$$= \boxed{\frac{1}{\sqrt{2\pi}}\left(\left(1-\frac{t}{\sqrt{n}}\right)^{\frac{t}{\sqrt{n}}+1}\left(1-\frac{t}{\sqrt{n}}\right)^{1-\frac{t}{\sqrt{n}}}\right)^{-n/2}}$$

F) let $x = +\frac{t}{\sqrt{n}}$

$$\frac{1}{\sqrt{2\pi}}\left(\left(1+x\right)^{1+x}\left(1-x\right)^{1-x}\right)^{-n/2}$$

$$= \frac{1}{\sqrt{2\pi}}\left(e^{(1+x)\ln(1+x)}\, e^{(1-x)\ln(1-x)}\right)^{-n/2}$$

$$= \frac{1}{\sqrt{2\pi}}\left(e^{\cancel{x}+\frac{x^2}{2}}\, e^{\cancel{-x}+\frac{x^2}{2}}\right)^{-n/2}$$

$$= \frac{1}{\sqrt{2\pi}}\left(e^{x^2}\right)^{-n/2} = \frac{1}{\sqrt{2\pi}}\left(e^{\frac{t^2}{n}}\right)^{-n/2} = \boxed{\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}}$$

3. a) binomial - m, p distribution with PMF

$P(i) = \binom{m}{i} p^i (1-p)^{m-i}$

given n samples $X_1 = x_1 \ldots X_n = x_n$

$$l(p) = P(X_1) \cdot P(X_2) \cdots P(X_n)$$

$$l(p) = \prod^n p(x_i)$$

$$\ln(l(p)) = \ln\left(\prod^n p(x_i)\right) = \sum^n \ln(p(x_i))$$

$$\frac{d}{dp} \ln(l(p)) = \frac{d}{dp} \sum^n \ln(p(x_i))$$

$$= \sum^n \frac{d}{dp} \ln(p(x_i))$$

$$= \sum^n \frac{d}{dp} \ln\left(\binom{m}{x_i} p^{x_i} (1-p)^{m-x_i}\right)$$

$$= \sum_i^n \frac{d}{dp}\left[\ln\binom{m}{x_i} + x_i \ln p + (m-x_i)\ln(1-p)\right]$$

$$= \sum_{i=1}^n \left(\frac{x_i}{p} - \frac{m-x_i}{1-p}\right) = 0$$

$$= \sum_{i=1}^n \left(\frac{x_i \cancel{- x_i p} - mp \cancel{+ x_i p}}{p(1-p)}\right) = 0$$

$$= \sum_{i=1}^n (x_i - mp) = 0$$

$$\Rightarrow nmp = \sum_{i=1}^n x_i$$

$$\boxed{p = \frac{1}{nm} \sum_{i=1}^n x_i}$$

b) $P(i) = (1-p)^{i-1} p \qquad i > 0$

given n samples $X_1 \ldots X_n$, max of $l(p)$?

$$p(l) = \prod^n p(x_i)$$

$$\ln(p(l)) = \ln\left(\prod^n p(x_i)\right) = \sum_i^n \ln(p(x_i))$$

$$0 = \frac{d}{dp} \ln(p(l)) = \frac{d}{dp} \sum^n \ln(p(x_i)) = \sum^n \frac{d}{dp} \ln((1-p)^{x_i - 1} p)$$

$$= \sum^n \frac{d}{dp}((x_i - 1)\ln(1-p) + \ln p)$$

$$= \sum_i^n \left(-\frac{x_i - 1}{1-p} + \frac{1}{p}\right) = 0$$

$$= \sum_i^n \frac{1 \cancel{-p} - x_i p \cancel{+p}}{p(1-p)} = 0$$

$$\sum_i^n (1 - x_i p) = 0 \longrightarrow n - p\sum^n x_i = 0$$

$$\boxed{p = \frac{n}{\sum_i^n x_i}}$$

c) $P(i) = \dfrac{\lambda^i e^{-\lambda}}{i!} \qquad i \geq 0$

$$\ell(\lambda) = \prod^n P(x_i)$$

$$\ln(\ell(\lambda)) = \sum^n \ln(P(x_i)) = \sum^n \ln\left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!}\right)$$

$$\frac{d}{dp}\ln(\ell(\lambda)) = \sum^n \frac{d}{dp}\ln\left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!}\right)$$

$$= \sum^n \frac{d}{dp}\left(x_i \ln\lambda - \lambda \ln e - \ln(x_i!)\right) = 0$$

$$= \sum^n \left(\frac{x_i}{\lambda} - 1\right) = 0$$

$$\frac{1}{\lambda}\sum^n x_i = n$$

$$\boxed{\lambda = \frac{\sum^n x_i}{n}}$$

4. $\hat{x} \equiv$ estimate $\quad X - \hat{x}$ = estimation error

a) $E[(X-\hat{x})^2] = E[X^2 - 2X\hat{x} + \hat{x}^2] = E[X^2] - 2\hat{x}E[X] + E[\hat{x}^2]$

$$\frac{dE[(X-\hat{x})^2]}{d\hat{x}} = 0 - 2E[X] + E[2\hat{x}] = 0$$

since derivatives are linear

$$= -2E[X] + 2E[\hat{x}] = 0$$

$$\Rightarrow 2E[X] = 2\hat{x}$$

$$\Rightarrow \boxed{\hat{x} = E[X]}$$

second deritive can show this is indeed min

So choosing $\hat{x}$ to be the expectation value minimizes the the mean squared error.

b) X, Y are random variables on same problem space

$$E[(X-\hat{x})^2] = E[(X-g(y))^2] = E[X^2 - 2g(y)X + g(y)^2]$$
$$= E[X^2] - 2g(y)E[X] + g(y)^2$$

We need to minimize the mean squared error with respect to the best function $g(y)$, so we can find a condition on $g(y)$ by taking the derivative of $E[(X-\hat{x})^2]$:

$$\frac{dE[(X-\hat{x})^2]}{d(g(y))} = 0 - 2E[X] + 2g(y)$$

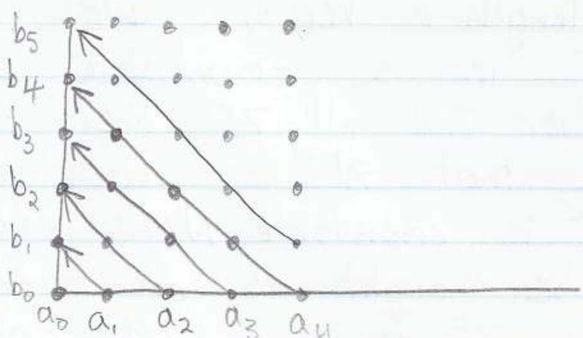
$$\Rightarrow \boxed{g(y) = E[X]}$$

since X and Y are independent, observing Y tells us virtually nothing about X, so our best estimate is just the same as part a.

5. S is <u>countable</u> if $\exists$ a bijection from N to S
S is countable if $\exists$ a surjection from N to S
S is countable if $\exists$ an injection from S to N

a) <u>Prove</u>: given 2 countable sets A & B, their cartesian product is countable.

Since A and B are countable, there is a mapping from each element in A and B to the natural numbers. So we can say that the element in A that maps to integer i is called a_i, and b_i is the element in B that maps to i. Then we can construct a grid like the following:



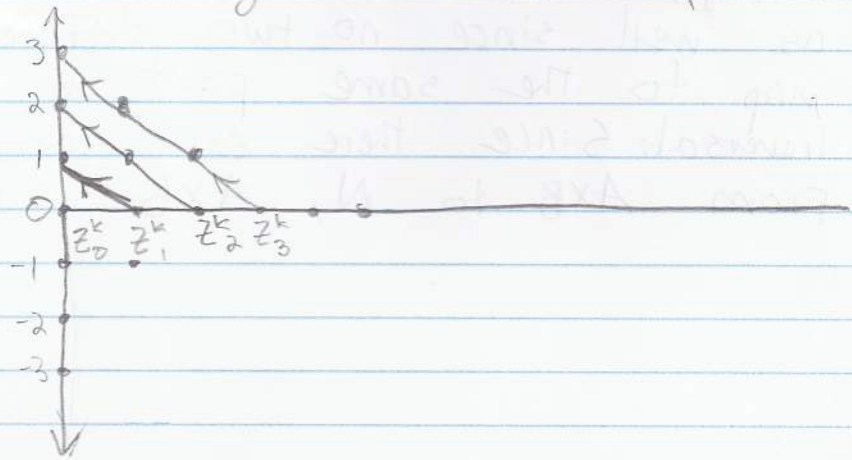
Then we can count <u>all</u> the elements in the set $A \times B$ by traversing the ordered pair diagonally. Notice that in this way we can in fact get all possible ordered pairs. The mapping from a point (a_i, b_j) to its natural number is an injection as well since no two different ordered pairs map to the same point on the diagonal traversal. Since there exist an injective mapping from $A \times B$ to N, $A \times B$ is countable.

b) Z^m set of all m-length vectors with integer elements For $m > 0$.

<u>Proof by Induction:</u>

base case
- $m=1$ $\quad Z^1$ = set of all length 1 vectors
- Since each length-1 vector just maps to the integer it contains, there is an injective mapping from Z^1 to N, so Z^1 is countable. (that is, no two different Z^1 vectors map to the same integer)

Assume true that for $m=k$ that Z^k is the set of all length k vectors with integer elements, and Z^k is a countable set.
To prove by induction that all Z^m are countable, we need to prove that Z^{k+1} is countable using the fact that Z^k is countable. Now every element in Z^{k+1} can just be thought as an ordered pair with an element in Z^k as its first coordinate, and an integer that is equal to the $(k+1)$th component for the second coordinate. Since each vector in Z^k is mappable to N, we can call a specific vector Z_i^k if its the ith vector, and set up an x-axis and y axis in an argument similar to part a):



we can count all the positive integer pairs diagonally as before, and assume that the negative pairs just double the total count, so the set of all $k+1$ vectors itself is countable. Using the diagonal traversal argument, we can hit every pairing of a vector from Z^k with an integer, so there is an injective mapping from Z^{k+1} to N. Thus Z^{k+1} is countable, and by induction, all Z^m are countable.

c) Prove: countable union of countable sets is countable
We can prove this by using a similar triangulated traversal to enumerate the union elements in such a way as to make an injective mapping. We can identify any element in the union by speficying which ith A_i set it came from and which jth element it in set A_i. Let's call this specified element A_{ij}. Even if the A_i's aren't disjoint, this proof still guarantees the mapping is injective since double counting an element in 2 different sets A_i still ensures that we count all union elements and no two different A_{ij}'s map to the same number. With 2 indices i and j, we can represent all elements in an $i \times j$ table, and diagonally enumerate all elements:

$$\begin{array}{ccccc}
A_{00} & A_{01} & A_{02} & A_{03} & \cdots \\
A_{10} & A_{11} & A_{12} & A_{13} & \cdots \\
A_{20} & A_{21} & A_{22} & A_{23} & \cdots \\
\vdots & \vdots & \vdots & &
\end{array}$$

Thus since all A_i are countable, i is either finite of countably infinite, and the number of unions is either finite or countably infinite. Our diagonal traversal shows we have an injection, so we have also proved that the union set itself is countable.

If there are countably infinite sets, then the union is also always infinite countable as well. We can proves this by looking at the simpliest case: if all A_i's only have 1 element, than we can enumerate all the elements in the union with the mapping $A_i \rightarrow i$. Thus we have directly constructed an injective mapping from the union to N, and we know The union set is countably infinite. Now any number of elements in each A_i would still make the union countably infinite. For example, if each A_i had 2 elements, then we could still make an injective mapping to N

6.a) No - it is impossible to write a program that takes n and returns the shortest way to represent n. This is similar to the compression problem we studied in class. Suppose n has m digits, so that there are 10^m strings of digits 0 through 9 that could possibly be n (we well let leading 0's be part of n - why not). Of course, any compression would have to be a string with $m-1$ characters. For the compression string that we return, we actually have a 15 letter alphabet (10 digits and [+, x, ^, (,)]), so there are 15^{m-1} possible compression strings that are actually shorter than n. But a large number of these strings don't actually evaluate to an integer, so in reality there are a lot fewer actual possible compression strings. Putting a lower bound on the number of uncompressible values of n:

$$\# \text{uncompressible} \geq 10^m - 15^{m-1}$$

we see that for any $m \geq 2$ this gives that the majority are uncompressible (you can't compress if n is only 1 character, a 0 character string wouldn't tell you much)

$m=2$: $10^2 - 15^1 = 100 - 15 = 85$

$m=3$ $10^3 - 15^2 = 1000 - 225 = 775$

Since most n's are uncompressible, we cannot write such a program as proposed.

8. Is the set of all finite length sequences of natural numbers countable? Are the algebraic numbers countable? Are the transcendental numbers uncountable? Are all numbers (complex and real) countable?

The set of all finite length sequences is a finitely countable union of length 1, length 2, ... sequences for all finite natural numbers. Since each set of certain finite length sequences is itself countable, we have the countable union of countable sets, and using our result from problem 5, we know that such a union is itself a countable set.

The algebraic numbers are defined as the numbers that are roots of polynomial equations. Using the argument, lets show that the polynomials are countable. The number of ternary strings is countable given that it is just a subset of the above set. Now each polynomial is uniquely defined by its coefficients, so we can convert it to a (degree d) d+1 vector of coefficients, which can be converted into binary and then converted to ternary, with 2's seperating each "binary" coefficient. Since every unique polynomial is injectively mapped to a ternary string, the polynomials are countable. Now every polynomial has a finite number of roots, so the number of algebraic roots is itself countable.

The transcendental numbers are all the numbers that are not algebraic. Since real numbers are uncountable, and algebraic numbers are countable, it follows that the Majority of real numbers are transcendental, and hence the transcendental numbers are uncountable.

The set of all numbers is uncountable, since it contains real numbers as a subset.