Solution to HW 5, Problem 8.3

Al-Naji, Nader

How long a string of random bits should be taken to be 50% sure that there are at least 32 consecutive 0's?

First, we compute the number of bitstrings of size N that have no runs of 32 consecutive zeros. We can find the generating function for this using Theorem 8.2. According to this, the generating function enumerating the number of bitstrings with no runs of p consecutive zeros is given by:

$$B_p(z) = \frac{1 - z^p}{1 - 2z + z^{p+1}}$$
$$\Rightarrow B_{32}(z) = \frac{1 - z^{32}}{1 - 2z + z^{33}}$$

Then, using Theorem 4.1, we can compute coefficients for an asymptotic. The following is the Mathematica code that implements Theorem 4.1 for numerator $f(z)$ and denominator $g(z)$:

$$\begin{aligned}
f[z_] &:= 1 - z^{32} \\
g[z_] &:= 1 - 2z + z^{33} \\
beta &:= 1/Min[Abs[z/.NSolve[g[z] == 0, z, 100]]] \\
v &:= 1 \\
c &:= -vbeta^v f[1/beta]/(D[g[z], z]//.z- > 1/beta)^v
\end{aligned}$$
$$NSolve[c(beta/2)^n == .5, 100]$$

In the last line, we use the approximation and divide by 2^n to get the proportion of strings of size N that have no runs of 32 zeros, which is also the probability. Plugging in:

$$f(z) = 1 - z^{32}$$
$$g(z) = 1 - 2z + z^{33}$$

and running the code yields:

$$\beta = 1.9999999997671693554787686282080979 2$$
$$c = 1.0000000034924596812625776675134911 9$$
$$c\left(\frac{\text{beta}}{2}\right)^N = 0.5$$
$$\Rightarrow N = 5.95409 \times 10^9$$

This is the number of bits we need to be 50% sure *no runs* of 32 consecutive zeros appears. Thus, it is also the number of bits needed to be 50% sure that *at least one run* of 32 consecutive zeros appears, since the events are complements.

Solution to HW 5, Problem 8.14

Al-Naji, Nader

Suppose that a monkey types randomly at a 32-key keyboard. What is the expected number of characters typed before the monkey hits upon the phrase "THE QUICK BROWN FOX JUMPED OVER THE LAZY DOG"?

Solution:
To do this problem, we extend Theorem 8.3 to M-character strings to find the generating function for the number of bitstrings not containing a pattern $p_0p_1...p_{P-1}$ where each character can be one of M values. We start with the OGF for S_p, the set of bitstrings with no occurrence of p:

$$\begin{aligned} S_p(z) &= \sum_{s \in S_p} z^{|s|} \\ &= \{\text{number of bitstrings of length N with no occurence of p}\} z^N \end{aligned}$$

Similarly, we define T_p to be the class of bitstrings that end with p but have no other occurrence of p, and name its associated generating function $T_p(z)$. Now, consider two symbolic relationships between S_p and T_p that translate to simulatneous equations involving $S_p(z)$ and $T_p(z)$. First, S_p and T_p are disjoint, and if we remove the last bit from a bitstring in either, we get a bitstring in S_p (or the empty bitstring). Expressed symbolically, this means that:

$$S_p + T_p = \epsilon + S_p \times (Z_0 + ... + Z_M)$$

which, since the OGF for $(Z_0 + ... + Z_M)$ is Mz, translates to:

$$S_p(z) + T_p(z) = 1 + MzS_p(z).$$

Second, consider the set of strings consisting of a string from S_p followed by the pattern. For each position i in the autocorrelation for the pattern, this gives a string from T_p followed by an $i - bit$ "tail". Expressed symbolically, this gives:

$$S_p \times < pattern >= T_p \times \sum_{c_i = 1} < tail >_i$$

which, since the OGF for $< pattern >$ is z^P and the OGF for $< tail >_i$ is z^i, translates to:

$$S_p(z) z^P = T_p(z) \sum_{c_i} z^i = T_p(z) c(z).$$

Then, since $c(z)$ is just one, we are left with the following two simultaneous equations:

$$\begin{aligned} S_p(z) + T_p(z) &= 1 + MzS_p(z) \\ S_p(z)z^P &= T_p(z) \end{aligned}$$

Solving this system yields:

$$\begin{aligned} S_p &= \frac{1}{z - M + z^P + 1} \\ T_p &= \frac{z^P}{z - M + z^P + 1} \end{aligned}$$

Now, we wish to compute:

$$\begin{aligned} &\sum_{N \geq 0} Pr(\text{index of end of first occurrence of } p \text{is greater than } N) \\ &= \sum_{N \geq 0} Pr(\text{first N characters do not contain p}) \\ &= \sum_{N \geq 0} \frac{\text{number of strings of length N that do not contain p}}{\text{total number of strings of length Ns}} \\ &= \sum_{N \geq 0} \frac{\text{number of strings of length N that do not contain p}}{M^N} \\ &= S_p(\frac{1}{M}) \\ S_p(\frac{1}{M}) &= \frac{1}{-M + (\frac{1}{M}) + \frac{1}{M}^P + 1} \\ S_p(\frac{1}{M}) &= M^P \\ &= 32^{44} \end{aligned}$$

Which gives us our final answer.

Solution to HW 5, Problem 8.57

Al-Naji, Nader

Solve the recurrence for p_n given in the proof of Theorem 8.9, to within the oscillating term:

$$p_n = \frac{1}{2^n}\sum_{0\le k\le n}\binom{n}{k}p_k \text{ for } n > 1 \text{ with } p_0 = 0 \text{ and } p_1 = 1.$$

Solution:
First, we define $P(z)$ to be an EGF with p_n as its coefficients:

$$P(z) = (0)\frac{z^0}{0!} + (1)\frac{z^1}{1!} + \sum_{N>1}\left(\frac{1}{2^n}\sum_{0\le k\le N}\binom{n}{k}p_k\right)\frac{z^n}{n!}$$

Then, in order to get a more concise expression, we compute $P(\frac{z}{2})e^{\frac{z}{2}}$. We do this using binomial convolution of the EGF for $e^{\frac{z}{2}}$ and $P(\frac{z}{2})$:

$$\begin{aligned}
P(\frac{z}{2}) &= (0)\frac{z^0}{2^0 0!} + (1)\frac{z^1}{2^1 1!} + \sum_{N>1}\left(\frac{1}{2^n}\sum_{0\le k\le N}\binom{n}{k}p_k\right)\frac{z^n}{2^n n!}\\
&= (0)\frac{z^0}{2^0 0!} + (1)\frac{z^1}{2^1 1!} + \sum_{N>1}\left(\frac{p_n}{2^n}\right)\frac{z^n}{n!}\\
e^{\frac{z}{2}} &= \sum_{n\ge 0}\frac{1}{2^n}\frac{z^n}{n!}\\
A(z)B(z) &= \sum_{n\ge 0}\left(\sum_{0\le k\le n}\binom{n}{k}a_k b_{n-k}\right)\frac{z^n}{n!}\\
\Rightarrow P(\frac{z}{2})e^{\frac{z}{2}} &= \sum_{n\ge 0}\left(\sum_{0\le k\le n}\binom{n}{k}\frac{p_k}{2^k}\frac{1}{2^{n-k}}\right)\frac{z^n}{n!}\\
&= \sum_{n\ge 0}\left(\frac{1}{2^n}\sum_{0\le k\le n}\binom{n}{k}p_k\right)\frac{z^n}{n!}\\
&= (0)\frac{z^0}{2^0 0!} + (1)\frac{z^1}{2^1 1!} + \sum_{n>1}\left(\frac{1}{2^n}\sum_{0\le k\le n}\binom{n}{k}p_k\right)\frac{z^n}{n!}
\end{aligned}$$

We now compute $P(z)$ in terms of the above function:

$$
\begin{aligned}
P(z) - P(\frac{z}{2})e^{\frac{z}{2}} &= \frac{z}{2} \\
\Rightarrow P(z) &= \frac{z}{2} + P(\frac{z}{2})e^{\frac{z}{2}} \\
\Rightarrow \frac{P(z)}{e^z} &= \frac{z}{2e^z} + \frac{P(\frac{z}{2})}{e^{\frac{z}{2}}}
\end{aligned}
$$

We can now unroll this recurrence:

$$
\begin{aligned}
\frac{P(z)}{e^z} &= \frac{z}{2e^z} + \frac{P(\frac{z}{2})}{e^{\frac{z}{2}}} \\
&= \frac{z}{2e^z} + \frac{z}{4e^{\frac{z}{2}}} + \frac{P(\frac{z}{4})}{e^{\frac{z}{4}}} \\
&= \frac{1}{2}\sum_{i>0} \frac{z}{2^i e^{\frac{z}{2^i}}} \\
\Rightarrow P(z) &= \frac{1}{2}\sum_{i>0} \frac{ze^z}{2^i e^{\frac{z}{2^i}}} \\
&= \frac{z}{2}\sum_{i>0} \frac{e^{z(1-\frac{1}{2^i})}}{2^i} \\
&= \frac{z}{2}\sum_{i>0} \frac{1}{2^i}\sum_{j\geq 0} \frac{(1-\frac{1}{2^i})^j z^j}{j!} \\
&= \frac{1}{2}\sum_{i>0} \frac{1}{2^i}\sum_{j\geq 0} \frac{(j+1)(1-\frac{1}{2^i})^j z^{j+1}}{(j+1)!} && \text{Multiplied by } \frac{j+1}{j+1} \\
&= \frac{1}{2}\sum_{i>0} \frac{1}{2^i}\sum_{j\geq 1} \frac{(j)(1-\frac{1}{2^i})^j z^j}{(j)!} && \text{Changed indices} \\
&= \sum_{j\geq 1}\sum_{i>0} \left(\frac{1}{2}\cdot\frac{1}{2^i}(j)(1-\frac{1}{2^i})^j\right)\frac{z^j}{j!} && \text{Changed indices} \\
\Rightarrow p_n &= \sum_{i>0}\left(\frac{1}{2}\cdot\frac{1}{2^i}n(1-\frac{1}{2^i})^n\right) \\
\Rightarrow p_n &= \sum_{i>0}\left(\frac{n}{2^{i+1}}(1-\frac{1}{2^i})^n\right) \\
&\sim \sum_{i>0}\left(\frac{n}{2^{i+1}}e^{\frac{1-n}{2^i}}\right)
\end{aligned}
$$

Now, we can use the Euler-Maclaurin Summation Formula on page 179, which states:

$$\sum_{a\leq k\leq b} f(k) = \int_a^b f(x)dx + \frac{f(a)+f(b)}{2} + \int_a^b (\{x\} - \frac{1}{2})f'(x)dx.$$

Letting $f(k) = \left(\frac{n}{2^{k+1}}e^{\frac{1-n}{2^k}}\right)$, we have:

$$= \int_a^b f(x)dx + \frac{f(a)+f(b)}{2} + \int_a^b (\{x\} - \frac{1}{2})f'(x)dx$$

At this point, I am lazy so I'm going to plug the expression into Mathematica and have it simplify everything down. After doing this, we get:

$$\frac{1}{4}e^{1-n}n + \frac{(e^n - e)\,n}{e^n(n-1)\log 4}$$
$$+ \int_0^\infty \left(-\frac{1}{2} + \{x\}\right)\left(-2^{-1-x}e^{2^{-x}(1-n)}n\log 2 - 2^{-1-2x}e^{2^{-x}(1-n)}(1-n)n\log 2\right)\,dx$$

This gives as our final answer:

$$p_n = \frac{1}{4}e^{1-n}n + \frac{(e^n - e)\,n}{e^n(n-1)\log 4} + \epsilon(n)$$

where $\epsilon(n) = \int_0^\infty \left(-\frac{1}{2} + \{x\}\right)\left(-2^{-1-x}e^{2^{-x}(1-n)}n\log 2 - 2^{-1-2x}e^{2^{-x}(1-n)}(1-n)n\log 2\right)\,dx$ is our oscillating term.

Solution to HW 5, Problem 9.3

Al-Naji, Nader

For $M = 365$, how many people are needed to be 99% sure that two have the same birthday?

Solution:
The probability that there are no collisions when N people have birthdays during an M-day calendar year is given by theorem 9.1:

$$(1-\frac{1}{M})(1-\frac{2}{M})...(1-\frac{N-1}{M}).$$

Since, we want the complement of this event to occur with probability 99%, we want to solve for this event occurring with probability 1%. We compute this below:

$$\begin{aligned}
(1-\frac{1}{M})(1-\frac{2}{M})...(1-\frac{N-1}{M}) &\sim \frac{1}{100} \\
\sum_{1\le k<N} \ln(1-\frac{k}{M}) &\sim \ln(\frac{1}{100}) \\
\sum_{1\le k<N} \frac{k}{M} &\sim \ln(100) \\
\frac{N(N-1)}{2M} &\sim \ln(200) \\
N &\sim \sqrt{2M\ln(100)}
\end{aligned}$$

Plugging in 365, we get $N \approx 57.9808$.

Solution to HW 5, Problem 9.38

Al-Naji, Nader

Prove that:

$$(\alpha+\beta)(n+\alpha+\beta)^{n-1} = \alpha\beta\sum_{k}\binom{n}{k}(k+\alpha)^{k-1}(n-k+\beta)^{n-k-1}$$

Solution:
We use the following lemma on page 528:

$$[z^n]g(C(z)) = \sum_{0\le k<n}(n-k)g_{n-k}\frac{n^{k-1}}{k!} \text{ when } g(z) = \sum_{k\ge 0} g_k z^k$$

Letting $g(z) = e^(\alpha z) = \sum_{k\ge 0}\frac{\alpha^k}{k!}z^k$, we use the theorem to compute $e^{\alpha C(z)}$:

$$\begin{aligned}
[z^n]g(C(z)) &= [z^n]e^{\alpha C(z)} \\
&= \sum_{0\le k<n}(n-k)\frac{\alpha^{n-k}}{(n-k)!}\frac{n^{k-1}}{k!} \\
&= \sum_{0\le k<n}\frac{\alpha^{n-k}}{(n-k-1)!}\frac{n^{k-1}}{k!} \\
&= \sum_{0\le k<n}\frac{\alpha^{n-k}}{(n-k-1)!}\frac{n^{k-1}}{k!} \\
&= \frac{\alpha}{n}\sum_{0\le k<n}\frac{1}{(n-1-k)!k!}n^k\alpha^{n-k-1} \\
&= \frac{\alpha}{n!}\sum_{0\le k<n}\frac{(n-1)!}{(n-1-k)!k!}n^k\alpha^{n-k-1} \\
&= \frac{\alpha}{n!}\sum_{0\le k<n}\binom{n-1}{k}n^k\alpha^{n-k-1} \\
&= \frac{\alpha}{n!}(\alpha+n)^{n-1}
\end{aligned}$$

Now, we use the fact that $e^{(\alpha+\beta)C(z)} = e^{\alpha C(z)}e^{\beta C(z)} \Rightarrow [z^n]e^{(\alpha+\beta)C(z)} = [z^n]e^{\alpha C(z)}e^{\beta C(z)}$ to complete the proof. Below, we use the binomial convolution identity from table 3.4:

$$\begin{aligned}
[z^n]e^{(\alpha+\beta)C(z)} &= \frac{\alpha+\beta}{n!}(\alpha+\beta+n)^{n-1} \\
[z^n]e^{\alpha C(z)}e^{\beta C(z)} &= \sum_{0\le k\le n} \frac{\binom{n}{k}a_k b_{n-k}}{n!} \\
&= \sum_{0\le k\le n} \frac{\binom{n}{k}\alpha(\alpha+k)^{k-1}\beta(\beta+n-k)^{n-k-1}}{n!} \\
&= \alpha\beta \sum_{0\le k\le n} \frac{\binom{n}{k}(\alpha+k)^{k-1}(\beta+n-k)^{n-k-1}}{n!} \\
\Rightarrow \frac{\alpha+\beta}{n!}(\alpha+\beta+n)^{n-1} &= \alpha\beta \sum_{0\le k\le n} \frac{\binom{n}{k}(\alpha+k)^{k-1}(\beta+n-k)^{n-k-1}}{n!} \\
\Rightarrow (\alpha+\beta)(\alpha+\beta+n)^{n-1} &= \alpha\beta \sum_{0\le k\le n} \binom{n}{k}(\alpha+k)^{k-1}(\beta+n-k)^{n-k-1}
\end{aligned}$$

Solution to HW 8, Problem 9.99

Al-Naji, Nader

Show that the probability that a random mapping of size N has no singleton cycles is 1/e, the same as for permutations (!).

Solution:
There are a total of N^N mappings of size N, since each position can map to N values. Further, there are $(N-1)^N$ mappings with no singleton cycles, since each position in a mapping with no singleton cycles can map to any value except itself (so $N-1$ values for each of the N positions). Thus, we have that the probability that a random mapping of size N has no singleton cycles is:

$$\begin{aligned}
\frac{(N-1)^N}{N^N} &= (\frac{N-1}{N})^N \\
&= (1-\frac{1}{N})^N \\
&= ((1+\frac{1}{-N})^{-N})^{-1} \\
&\sim (e)^{-1} \\
&\sim \frac{1}{e}
\end{aligned}$$