1. You toss a fair coin until you get head twice. What is the probability that you made n tosses?

The function for the Fibonacci Numbers is $\dfrac{x}{1-x-x^2}$

Thus, the number of ways to end at k flips is F_{k-1} so the probability of ending at k flips is $\dfrac{F_{k-1}}{2^k}$

2. Each person in a group of n people is requested to select a number between 1 to k. Describe the probability that at least 2 people chose the same number.

There are k choices and n people.

- First person will choose a random number.
- Second person will choose a random number and will have $\frac{1}{k}$ probability to choose the number chosen by others.
- Third person will have $\frac{2}{k}$ probability.
- n-th person will have $\frac{n-1}{k}$ probability.

Geting the probability that way will be way too hard, let's get the probability of *not* having two identical choices:

- First person will choose a random number (and have $\frac{k}{k}$ probability).
- Second person will choose a random number and have $\frac{k-1}{k}$ to choose the number not chosen by others.
- Third person will have $\frac{k-2}{k}$ probability.
- n-th person will have $\frac{k-n+1}{k}$ probability.

Probability (p) is then:

$$p = 1 - \frac{k(k-1)(k-2)(k-3)\ldots(k-n+3)(k-n+2)(k-n+1)}{k^n} = 1 - \frac{k!}{k^n(k-n)!}$$

3. The frequency of the Malum Indentum disease in the population is 1 in 10,000. A test that checks if one is infected with the disease is 99% accurate. One takes the test and gets a positive response (test says she is infected). What's the probability that she is infected?

we can use Baise theorem here

$P(A|B) = (P(A) * P(B|A)) / P(B)$

A - ill $\qquad$ $\overline{A}$ - is not ill

B - test positive

$$P(A|B) = \frac{P(A|B) * P(A)}{P(B|A) * P(\overline{A}) + P(B|\overline{A}) * P(\overline{A})}$$

$P(A) = 0{,}001$

$P(B|A) = 0{,}99$

$P(\overline{A}) = 0{,}9999$

$P(B|\overline{A}) = 0{,}01$

$$P(A|B) = \frac{0{,}99 * 0{,}0001}{0{,}99 * 0{,}0001 + 0{,}01 * 0{,}9999} = 0{,}0098$$

4. Let X and Y be discrete random variables, Z be a continuous random variable, and α and β constants.

Prove the following qualities:

a. E(X + Y) = E(X) + E(Y) ,

Then we can consider the random variable $X + Y$ to be the result of applying the function $\varphi(x, y) = x+y$ to the joint random variable (X, Y).

$$\begin{aligned} E(X+Y) &= \sum_j \sum_k (x_j + y_k) P(X = x_j,\ Y = y_k) \\ &= \sum_j \sum_k x_j P(X = x_j,\ Y = y_k) + \sum_j \sum_k y_k P(X = x_j,\ Y = y_k) \\ &= \sum_j x_j P(X = x_j) + \sum_k y_k P(Y = y_k)\ . \end{aligned}$$

The last equality follows from the fact that

$$\sum_k P(X = x_j,\ Y = y_k) = P(X = x_j)$$

and

$$\sum_j P(X = x_j,\ Y = y_k) = P(Y = y_k)\ .$$

Thus, $E(X + Y) = E(X) + E(Y)$

b. E(αZ) = αE(Z)

$$\begin{aligned} E(\alpha Z) &= \sum_j \alpha z_j P(Z = z_j) \\ &= \alpha \sum_j z_j P(Z = z_j) \\ &= \alpha E(Z) \ . \end{aligned}$$

c. If X and Y are independent then E(XY) = E(X)E(Y)

Where X and Y are continuous random variables, by definition they are independent when $f_{XY}(x, y) = f_X(x) f_Y(y)$. Then we have

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty} f_{XY}(x, y) xy dx dy \\ &= \int_{-\infty}^{\infty} f_X(x) f_Y(y) xy dx dy \\ &= \int_{-\infty}^{\infty} f_X(x) x dx \int_{-\infty}^{\infty} f_Y(y) y dy \\ &= E(X)E(Y) \end{aligned}$$

d. $V(\alpha X + \beta) = \alpha^2 V(X)$

$$\begin{aligned} V(\alpha X + \beta) &= E\left[((\alpha X + \beta) - E[\alpha X + \beta])^2\right] \\ &= E\left[(\alpha^2 X + \beta - \alpha E[X] - \beta)^2\right] \\ &= E\left[\alpha^2 (X - E[X])^2\right] \\ &= \alpha^2 E\left[(X - E[X])^2\right] = \alpha^2 V(X) \end{aligned}$$

e. If X and Y are independent then V(X+Y) = V(X)+V(Y)

$$\begin{aligned} V(X+Y) &= E((X+Y)^2) - (E(X+Y))^2 \\ &= E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2 \\ &= E(X^2) + 2E(X)E(Y) + E(Y^2) - (E(X))^2 - 2E(X)E(Y) - (E(Y))^2 \\ &\qquad \text{using } E(XY) = E(X)E(Y) \text{ at the start since } X, Y \text{ independent} \\ &= E(X^2) - (E(X))^2 + E(Y^2) - (E(Y))^2 \\ &= V(X) + V(Y) \end{aligned}$$

5. Let X_i ~ Unif(0, 1) for 1 <= i <= n be IID (independent identically distributed) random variables.

   Let Y = max(X_1, ..., X_n). What is E(Y)?

$$
\begin{aligned}
E[Y] &= \int_0^1 \cdots \int_0^1 \max(x_1, \ldots, x_n) dx_1 \ldots dx_n \\
&= n! \int_{x_1 < \ldots < x_n} \max(x_1, \ldots, x_n) dx_1 \ldots dx_n \\
&\quad \text{since the integral is the same over all } n! \text{ such regions} \\
&= n! \int_{x_1 < \ldots < x_n} x_n \, dx_1 \ldots dx_n \\
&\quad \text{since the max in this region is } x_n \\
&= n! \int_0^1 \int_0^{x_n} \int_0^{x_{n-1}} \cdots \int_0^{x_2} x_n \, dx_1 \ldots dx_n \\
&= n! \int_0^1 \int_0^{x_n} \int_0^{x_{n-1}} \cdots \int_0^{x_3} x_n x_2 \, dx_2 \ldots dx_n \\
&= n! \int_0^1 \int_0^{x_n} \int_0^{x_{n-1}} \cdots \int_0^{x_4} x_n \, \frac{1}{2} x_3^2 \, dx_3 \ldots dx_n \\
&= \ldots \\
&= n! \int_0^1 x_n \cdot \frac{1}{(n-1)!} x_n^{n-1} \, dx_n \\
&= n \int_0^1 x_n^n \, dx_n \\
&= \frac{n}{n+1}.
\end{aligned}
$$

6. A drunken point hops on the number line, making jumps sized 1. The probability to jump to the right is fixed: P(right) = p. Let X_n be the position of the point after n jumps.

   a. What is $E(X_n)$?

   b. What is $V(X_n)$?

P(right) = p
P(left) = 1-p

if we do k steps to right we also do (n-k) steps to left

k - (n-k) = 2k - n = m <-- X_n - the position

$$k = \frac{m+n}{2}$$

so we need m + n to be even
P_1 = 0 if m + n is an odd number

$$P = C_n^{\frac{m+n}{2}} * (p)^{\frac{m+n}{2}} * (1-p)^{\frac{m+n}{2}}$$

According Bernaulli distribution

$$E(X) = \sum_{k=0} 2(k-n) * C_n^k * p^k * (1-p)^{n-k}$$

$$V(x) = E(X^2) - (E(X))^2$$