Penguin anomaly

Excerpt from *'Penguin' Anomaly Hints at Missing Particles.*[1]

> Most unexpected bumps in data go away as more data accumulate, just as you might get seven heads in your first 10 coin tosses only to end up with a 50-50 ratio after many more tosses. But after tripling their original sample size and analyzing approximately 2,400 of the rare penguin decays, the LHCb scientists say the anomaly hasn't diminished. Instead, it has lingered at an estimated statistical significance of "3.7 sigma" which means it is just as unlikely for such a large fluctuation to happen randomly as it would be to get 69 heads in 100 coin tosses. Physicists require a 5-sigma deviation from their expectations, equivalent to flipping 75 heads in 100 tosses (the odds of which are less than one in a million), to claim the discovery of a real effect.

Recall that the binomial mass function with $p = 1/2$ is the probability of obtaining exactly k heads in n tosses.

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

In Eigenmath, define the binomial mass function and calculate the probability of getting exactly 75 heads in 100 tosses.

```
f(k) = choose(n,k) p^k (1-p)^(n-k)
n = 100
p = 1/2
float(f(75))
```

1.91314×10^{-7}

Hence the probability of getting exactly 75 heads in 100 tosses is indeed less than one in a million.

As the following equation shows, the variance for 100 coin tosses is 25.

$$\sigma^2 = np(1-p) = 100 \times \tfrac{1}{2} \times \tfrac{1}{2} = 25$$

Hence $\sigma = 5$. It follows that a 5-sigma deviation from the expected value is $50 + 5\sigma = 75$ and a 3.7-sigma deviation is $50 + 3.7\sigma = 68.5$.

Flipping 75 heads is a rare event, but flipping 25 or 76 is also rare. Let us compute the probability of all rare events. Start with the following cumulative distribution function.

$$\Pr(X \le x) = F(x) = \sum_{k=0}^{x} f(k)$$

[1] Wolchover, Natalie. *'Penguin' Anomaly Hints at Missing Particles.* www.quantamagazine.org/penguin-anomaly-at-large-hadron-collider-hints-at-missing-particles-20150320/

Define a “rare” result as a total number of heads that is beyond 5 standard deviations of the mean $\mu = 50$. Then the probability of a “rare” result after 100 tosses is

$$\Pr(X \le 25) + \Pr(X \ge 75) = F(25) + 1 - F(74) = 2\,F(25)$$

The last equivalence is by symmetry of the binomial function.

In Eigenmath, define the cumulative distribution function and compute the probability of fewer than 26 or more than 74 heads in 100 tosses.

```
F(x) = sum(k,0,x,f(k))
float(2 F(25))
```

5.63628×10^{-7}

The probability is still less than one in a million.