The balancing feature of the tPrimes case generator augments the frequency of True/False such that the frequency of the output is equal to the argument passed. This is especially important when the sample size is large and the exponents are not.

Variables and Conditions

a: Lower bound on exponents

b: Upper bound on exponents

j: Number of sample groups d: Exponent difference between sample groups

X: Random variable denoting the solution should be true or false

$X \subseteq \{0, 1\}$

K: Random variable denoting the exponent range

$K \subseteq \{[a, a + d], [a + d, a + 2d], \dots, [b - d, b]\}$

$$P(X = 1) = \sum_{k=1}^{j} P_k(X = 1)P(K = k)$$

We can split $\sum_{k=1}^{j} P_k(X = 1)P(K = k)$ into two groups: complete and incomplete–demarcated by h

$$P(X = 1) = \sum_{k=1}^{h} P_k(X = 1)P(K = k) + \sum_{k=h}^{j} P_k(X = 1)P(K = k)$$

Subtract the completed

$$\sum_{k=h}^{j} P_k(X = 1)P(K = k) = P(X = 1) - \sum_{k=1}^{h} P_k(X = 1)P(K = k)$$

$P(K = k)$ is uniform for all ranges. Denote this shared probability as p and divide all by it

$$\sum_{k=h}^{j} P_k(X=1) = \frac{P(X=1)}{p} - \sum_{k=1}^{h} P_k(X=1)$$

If we want to re-adjust the average probability for incomplete samples, divide by the remaining number of groups

$$\frac{\sum_{k=h}^{j} P_k(X=1)}{j-h} = \frac{\frac{P(X=1)}{p} - \sum_{k=1}^{h} P_k(X=1)}{j-h}$$

Denote the new average probability for all samples going forward as μ_h

$$\mu_h = \frac{\sum_{k=h}^{j} P_k(X=1)}{j-h}$$

By association

$$\mu_h = \frac{\frac{P(X=1)}{p} - \sum_{k=1}^{h} P_k(X=1)}{j-h}$$

Because this is uniform

$$p = \frac{1}{j}$$

$$\mu_h = \frac{P(X=1)}{\frac{1}{j}(j-h)} - \frac{\sum_{k=1}^{h} P_k(X=1)}{j-h}$$

$$\mu_h = \frac{P(X=1)}{(1-\frac{h}{j})} - \frac{\sum_{k=1}^{h} P_k(X=1)}{j-h}$$