CURRICULUM, PEDAGOGY AND BEYOND

THE MATHEMATICAL ASSOCIATION OF VICTORIA

MAV24 CONFERENCE

THE MATHEMATICAL ASSOCIATION OF VICTORIA

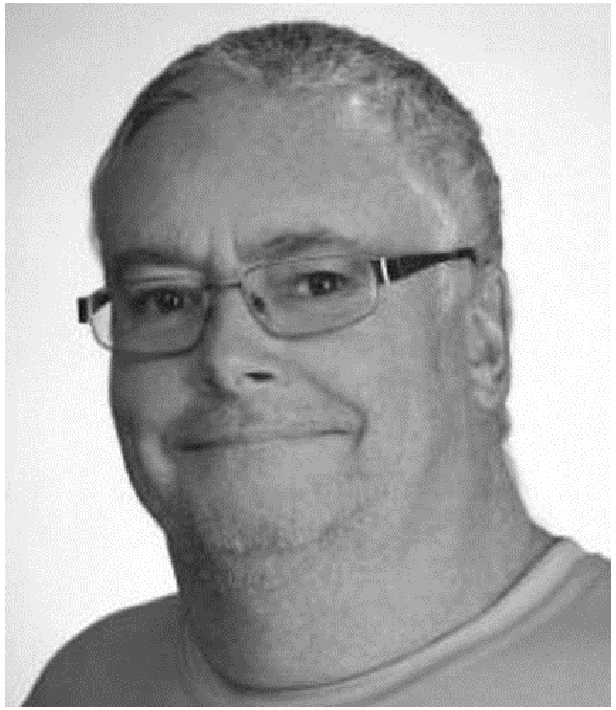
MAV24 CONFERENCE

Card Collector: An interesting problem for investigation.

Brian Lannen

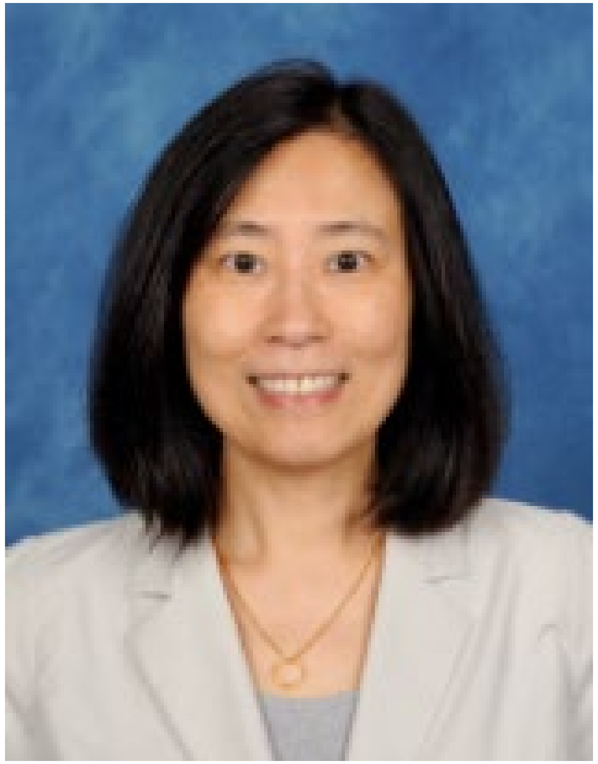
Murray Mathematics Curriculum Services

Brian Lannen has been teaching for over 35 years in Victoria and NSW. He has taught Physics, Maths and Science in schools, university and TAFE colleges, was a curriculum consultant in NSW and New York and has contributed to a range of text-book writing projects. He helped establish T-Cubed (Teachers Teaching with Technology) in Australia in the 1990s and is now a Senior Mentor to that association and Principal Host of the Texas Instruments Australia webinar program.



Angel Wong

St. Andrew’s Christian College



With over 15 years of classroom experience, I have taught Mathematics, Science, and Biblical Studies, bringing a rich, interdisciplinary perspective to education. Formerly a computer engineer, I developed a deep appreciation for technology's role in solving complex problems, a passion that now drives my teaching philosophy. Committed to inspiring creativity and fostering innovative thinking, I integrate my extensive knowledge of Mathematics and technology to equip students with essential skills for future challenges. Enthusiastic about merging technology with Mathematics education, I promote a learning environment that encourages curiosity and critical thinking. In my upcoming presentation at the Maths Association Victoria, I will showcase how technology can inspire creativity and innovation in the Mathematics classroom.

Session Description

- Consider that a breakfast cereal company is running a promotion by inserting a famous mathematician card into each box of cereal. If there are 6 different cards in the set and the placement of cards is equal and random, how many boxes of cereal *on average* would you expect are needed in order to collect the full set of cards?
- Participants will partake in hands-on investigation of the card (or coupon) collector problem with the opportunity to explore through intuition, simulation and calculation. This classic problem could be used as the basis for a Mathematical Methods Investigation task, specifically examining expectation, probability distribution and confidence intervals. However, at a basic level, it is also accessible to junior students and can be used to meet the F-10 curriculum requirement to "conduct simulations, using digital tools to determine probabilities and describe results (VC2M8P03)."

Victorian Certificate of Educatio

Mathematics

Study Design

Accreditation Period

2023–2027

Key skills

- use computational thinking, algorithms, models and simulations to solve problems related to a given context
- simulation using simple random generators such as coins, dice, spinners and pseudo-random generators using technology, and the display and interpretation of results, including informal consideration of proportions in samples
- set up probability simulations, and describe the notions of randomness and variability, and their relation to events
- the purpose and effect of sequencing, decision-making and repetition statements on relevant functionalities of technology, and their role in the design of algorithms and simulations
- simulation to estimate probabilities involving selection with and without replacement.

Simulation

- random experiments, events and event spaces
- use of simulation to generate a random sample.

Investigation using simulation in Victorian F-10 Maths curriculum V. 2.0

Victorian Curriculum
Foundation–10

Mathematics Version 2.0

Probability content descriptions include:

Level 7 conduct repeated chance experiments and run simulations with a **large number of trials** using digital tools; **compare** predicted with observed results (VC2M7P02)

Level 8 conduct repeated chance experiments and simulations, using digital tools to determine probabilities for **compound events**, and describe results (VC2M8P03)

Level 9 **design** and conduct repeated chance experiments and simulations using digital tools to **estimate** probabilities that cannot be determined exactly (VC2M9P03)

Level 10 ... design and conduct simulations using digital tools to model **conditional** probability and **interpret** results (VC2M10P01)

Card (or Coupon) Collector Problem

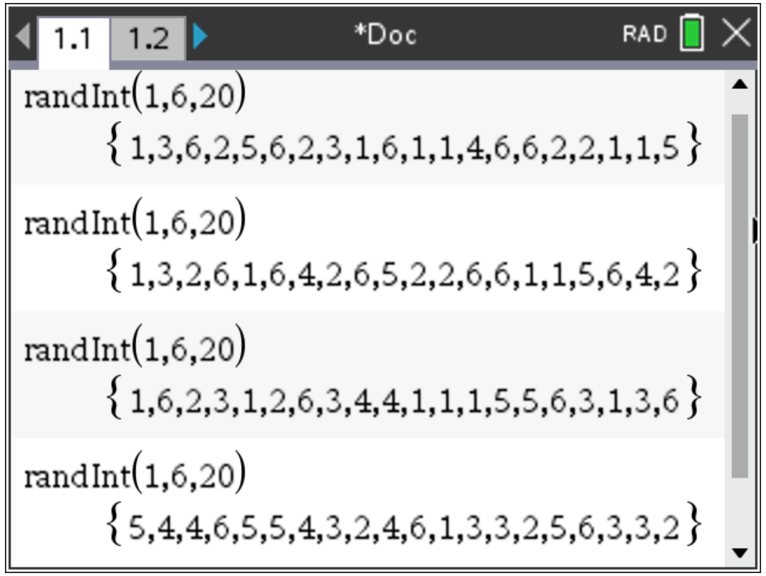
Consider that a breakfast cereal company is running a promotion by inserting a famous footballer card into each box of cereal. If there are 6 different cards in the set and the placement of cards is equal and random, how many boxes of cereal *on average* would you expect are needed in order to collect the full set of cards?

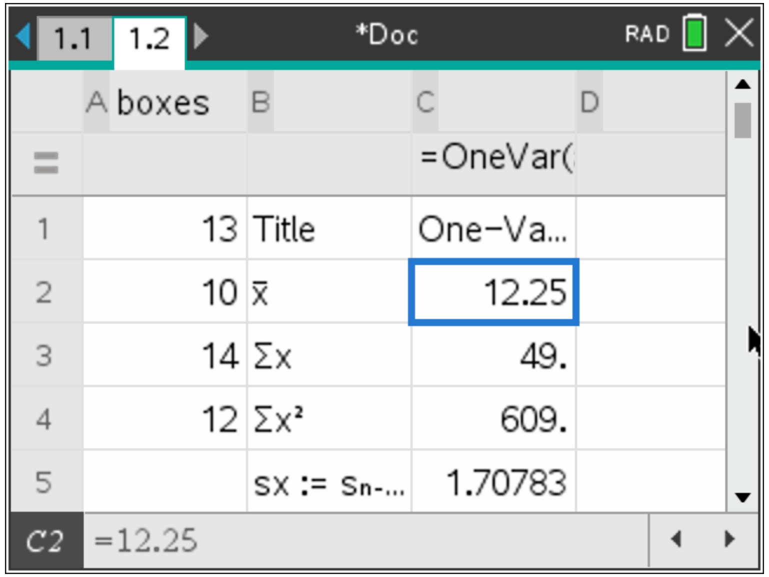
Card (or Coupon) Collector Problem

Consider that a breakfast cereal company is running a promotion by inserting a famous **female mathematician** card into each box of cereal. If there are 6 different cards in the set and the placement of cards is equal and random, how many boxes of cereal *on average* would you expect are needed in order to collect the full set of cards?

Intuition → Simulation → Calculation → Investigation

Electronic Simulation





Background knowledge

Bernoulli Distribution

- A discrete probability distribution.
- Models a single trial with two possible outcomes:
  Success (1)
  Failure (0)
- Example: Coin flip (heads = 1, tails = 0).
- Product quality test (Pass = 1, Fail = 0)
- Requirement (Met = 1, Not Met = 0)

Background knowledge

Bernoulli Distribution

- Probability of Success (p)
- Probability of Failure (1 – p)
- Random Variable (X)

x	0	1
Pr (X = x)	1-p	p

- E(X) = $\sum xp(x)$ = 0 (1-p) + 1 (p) = p
- Var(X) = $\sum (x-\mu)^2 p(x)$ = p (1-p)

Background knowledge

Geometric Distribution

- A discrete probability distribution modelling the number of Trials needed to achieve the first success

- Each trial is independent with:
  - Probability of success p
  - Probability of failure $1 - p$

- Example: Rolling a die until you get a 6

Background knowledge

Geometric Distribution

- For the first success to be at *k*-th trial:
  - First (k -1) trails are failures
  - The k-th trial is success
- Probability for this event is :

$$\begin{aligned} P(X=k) &= \underbrace{(1-p)\times\ldots\times(1-p)}_{(k-1)\ \text{times}}\times p \\ &= (1-p)^{k-1}\,p. \end{aligned}$$

Background knowledge

Geometric Distribution

- Random Variable (X) – Number of trials until the first success
- Success Probability (p) – Constant probability of success in
- each trial

x	1	2	3	4	...	k
Pr (X = x)	P	(1-p)(p)	$(1-p)^2(p)$	$(1-p)^3(p)$		$(1-p)^{k-1}(p)$

$$E[X] = p + \dots + k(1-p)^{k-1}p + \dots$$

$$= p\sum_{k=1}^{\infty} k(1-p)^{k-1} = \frac{1}{p}.$$

Back to coupon collect problem

Repeated Geometric Distribution

To get the k-th coupon, after collecting $k - 1$ coupons:

Failure = getting coupon you already have

There are $k - 1$ choices

Success = getting a new coupon.

There are $n - k + 1$, where n is the total number of coupons to be collected

Example: If there are 6 (1, 2,..6) coupons to be collected and we have collected cards 1, 3, 4:

$n = 6$, $k = 4$ (to get the 4^{th} coupon)

$k - 1 = 3$

$n - (k - 1) = n - k + 1 = 3$

Back to coupon collect problem

T_k is the discrete random variable representing number of purchases needed to get the *k*-th coupons after collecting *k-1* coupons

T_k is geometric random variable with success probability $p = \frac{n-k+1}{n}$

Therefore $E(T_k) = \frac{1}{p} = \frac{n}{n-k+1}$

Calculation

Consider that a breakfast cereal company is running a promotion by inserting a card into each box of cereal. If there are 6 different cards in the set and the placement of cards is equal and random, how many boxes of cereal *on average* would you **expect** are needed in order to collect the full set of cards?

E(X)

where X = number of cereal boxes needed to collect a set of N unique cards.

Calculation

Hoping for a Bernoulli trial "success" by adding the k^{th} discrete card to my collection on the opening of next box.	k						

Calculation

Hoping for a Bernoulli trial "success" by adding the k^{th} discrete card to my collection on the opening of next box.	k	1	2	3	4	5	6

Calculation

$$Pr(k) = \frac{N-(k-1)}{N} \quad E(X_k) = \frac{1}{Pr(k)}$$

Number of unique cards I already have	k-1	0	1	2	3	4	5
Hoping for a Bernoulli trial "success" by adding the k^{th} discrete card to my collection on the opening of next box.	k	1	2	3	4	5	6

Calculation

$$Pr(k) = \frac{N-(k-1)}{N} \qquad E(X_k) = \frac{1}{Pr(k)}$$

Number of unique cards I already have	k-1	0	1	2	3	4	5
Expected number of new cereal boxes needed in order to obtain the k^{th} discrete card.	$\frac{N}{N-k+1}$	$\frac{N}{N}$	$\frac{N}{N-1}$	$\frac{N}{N-2}$	$\frac{N}{N-3}$	$\frac{N}{N-4}$	$\frac{N}{N-5}$

Calculation

$$Pr(k) = \frac{N-(k-1)}{N} \qquad E(X_k) = \frac{1}{Pr(k)}$$

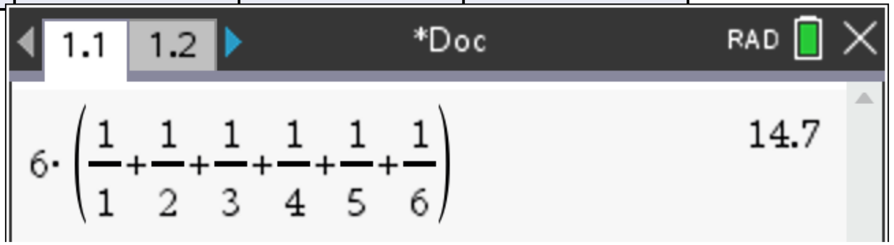
Number of unique cards I already have	k-1	0	1	2	3	4	5
Expected number of new cereal boxes needed in order to obtain the k^{th} discrete card.	$\frac{N}{N-k+1}$	$\frac{N}{N}$	$\frac{N}{N-1}$	$\frac{N}{N-2}$	$\frac{N}{N-3}$	$\frac{N}{N-4}$	$\frac{N}{N-5}$
Example when N=6	$E(X_k)$						

Calculation

$$Pr(k) = \frac{N-(k-1)}{N} \qquad E(X_k) = \frac{1}{Pr(k)}$$

Number of unique cards I already have	k-1	0	1	2	3	4	5
Expected number of new cereal boxes needed in order to obtain the kth discrete card.	$\frac{N}{N-k+1}$	$\frac{N}{N}$	$\frac{N}{N-1}$	$\frac{N}{N-2}$	$\frac{N}{N-3}$	$\frac{N}{N-4}$	$\frac{N}{N-5}$
Example when N=6	$E(X_k)$	$\frac{6}{6}$	$\frac{6}{5}$	$\frac{6}{4}$	$\frac{6}{3}$	$\frac{6}{2}$	$\frac{6}{1}$

$$E(X) = \frac{6}{6} + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1} = \frac{147}{10} = 14.7$$



Calculation

1.1 | 1.2 | *Doc | RAD

$$6\cdot\left(\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}\right) \quad 14.7$$

$$6\cdot\sum_{i=1}^{6}\left(\frac{1}{i}\right) \quad 14.7$$

$$12\cdot\sum_{i=1}^{12}\left(\frac{1}{i}\right) \quad 37.2385$$

$$\mathrm{E}[X] = nH_n$$

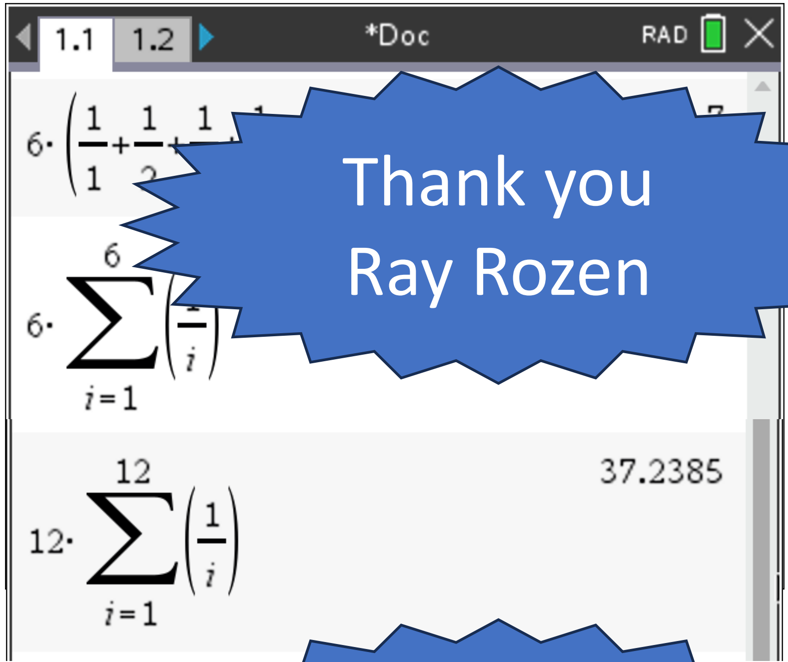
Where H_n is the n^{th} harmonic number.

$$\mathrm{E}[X] = \sum_{k=1}^{n}\mathrm{E}[X_k]$$

$$= n\sum_{k=1}^{n}\frac{1}{k}$$

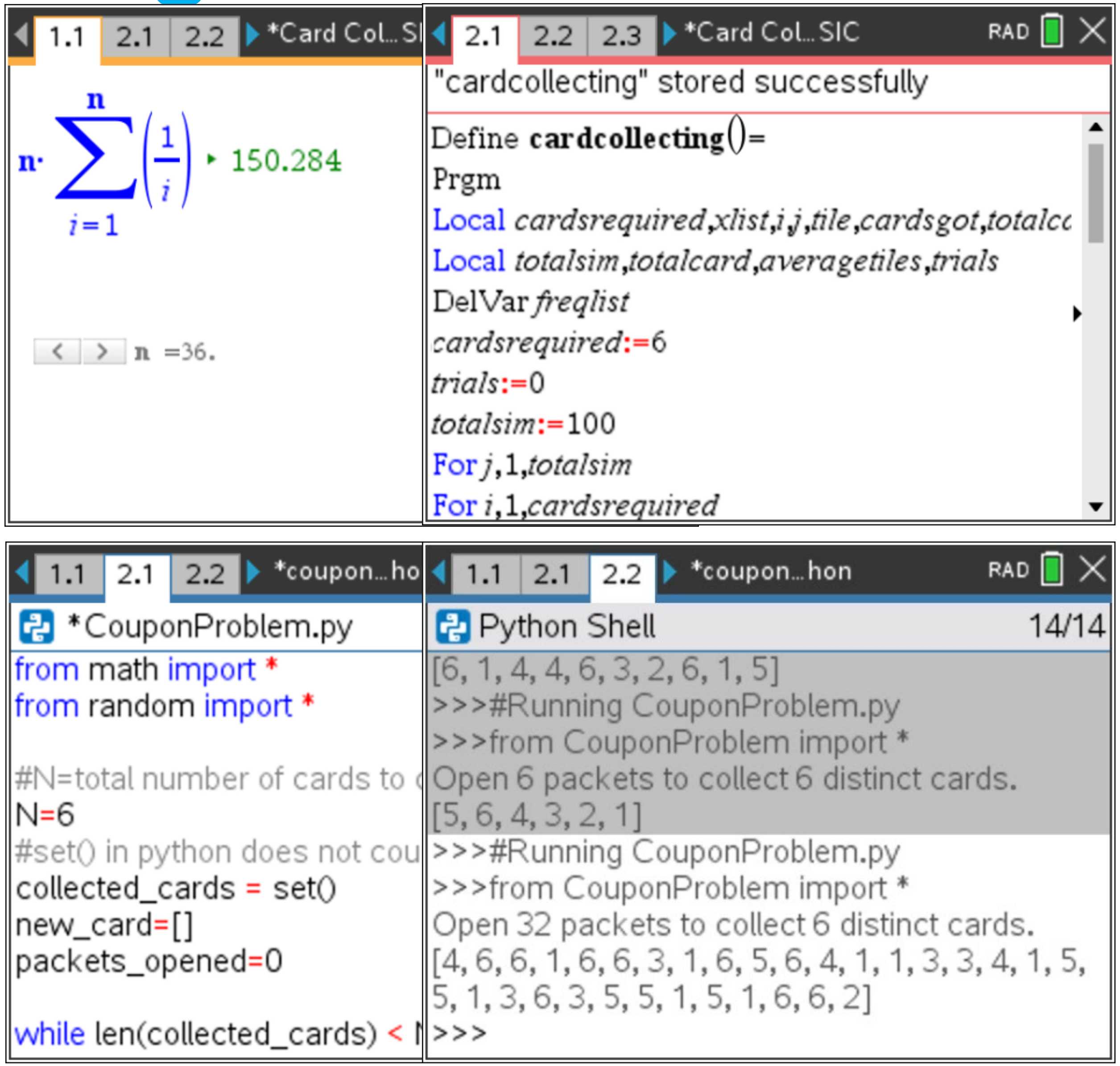
Calculation

Programs

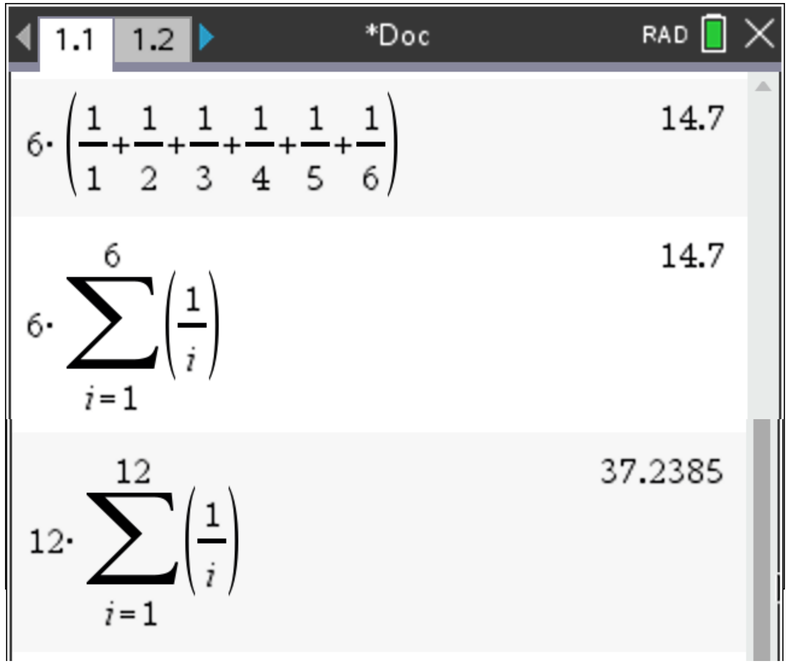


Thank you
Ray Rozen

Thank you
Angel Wong



Calculation

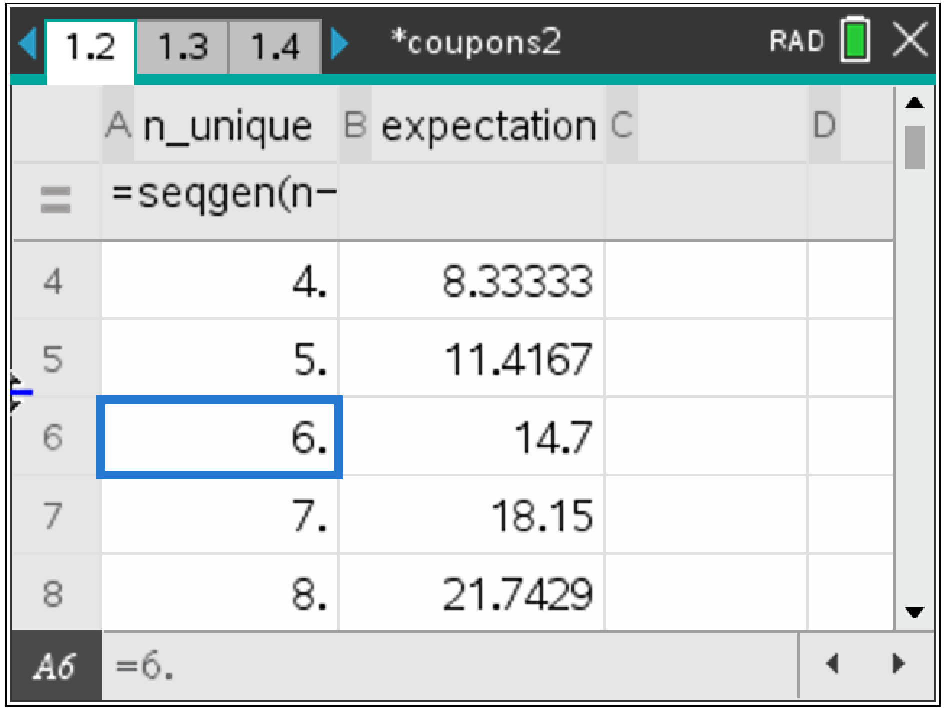


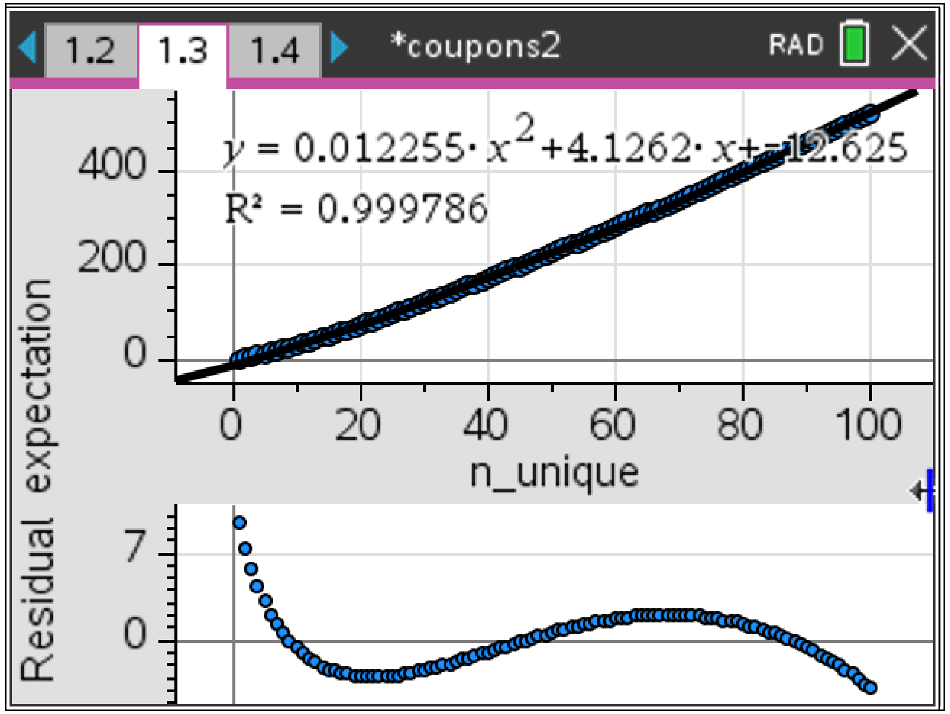
Investigation

- How does the expectation relate to the number of different cards in the set?
- Can we also calculate the standard deviation?
- What does the distribution look like?
- What if the cards were not of equal likelihood?

Investigation

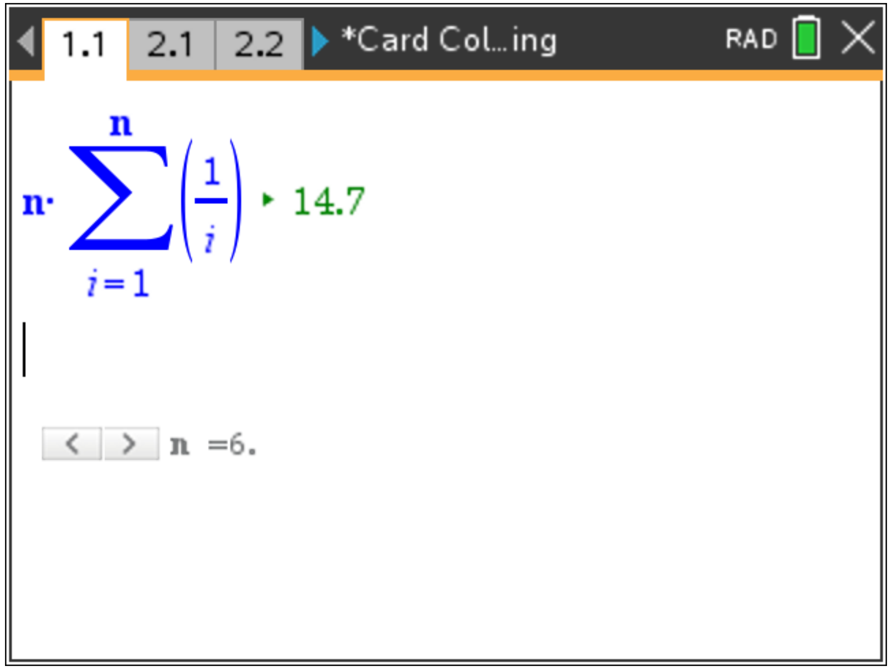
- How does the expectation relate to the number of different cards in the set?

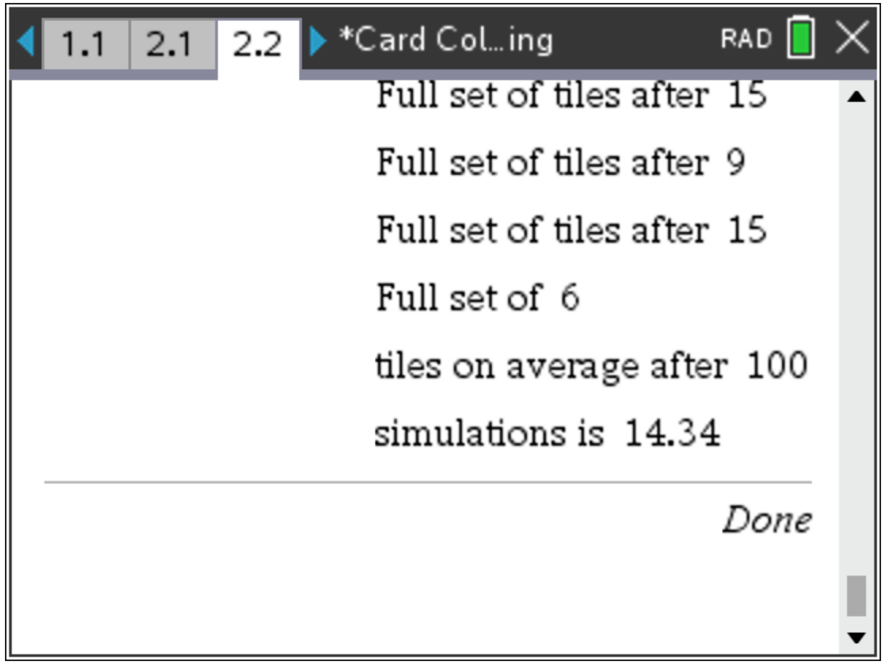




Investigation

- How does the expectation relate to the number of different cards in the set?





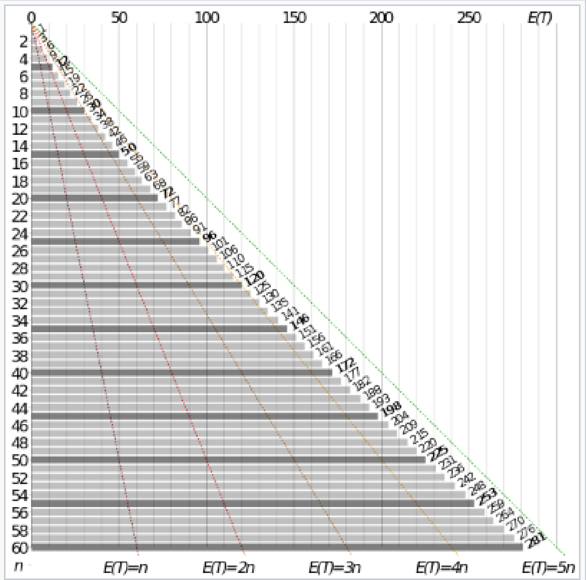
Coupon collector's problem

14 languages

Article Talk Read Edit View history Tools

From Wikipedia, the free encyclopedia

In probability theory, the **coupon collector's problem** describes "collect all coupons and win" contests. It asks the following question: If each box of a brand of cereals contains a coupon, and there are *n* different types of coupons, what is the probability that more than *t* boxes need to be bought to collect all *n* coupons? An alternative statement is: Given *n* coupons, how many coupons do you expect you need to draw with replacement before having drawn each coupon at least once? The mathematical analysis of the problem reveals that the expected number of trials needed grows as $\Theta(n\log(n))$.[a] For example, when $n = 50$ it takes about 225[b] trials on average to collect all 50 coupons.



Graph of number of coupons, *n* vs the expected number of trials (i.e., time) needed to collect them all, *E* (*T*)

Solution [edit]

Calculating the expectation [edit]

Let time *T* be the number of draws needed to collect all *n* coupons, and let t_i be the time to collect the *i*-th coupon after $i - 1$ coupons have been collected. Then $T = t_1 + \cdots + t_n$. Think of *T* and t_i as random variables. Observe that the probability of collecting a *new* coupon is $p_i = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$. Therefore, t_i has geometric distribution with expectation $\frac{1}{p_i} = \frac{n}{n-i+1}$. By the linearity of expectations we have:

$$\begin{aligned}
\operatorname{E}(T) &= \operatorname{E}(t_1 + t_2 + \cdots + t_n) \\
&= \operatorname{E}(t_1) + \operatorname{E}(t_2) + \cdots + \operatorname{E}(t_n) \\
&= \frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n} \\
&= \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{1} \\
&= n \cdot \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n}\right) \\
&= n \cdot H_n.
\end{aligned}$$

Here H_n is the *n*-th harmonic number. Using the asymptotics of the harmonic numbers, we obtain:

$$\operatorname{E}(T) = n \cdot H_n = n\log n + \gamma n + \frac{1}{2} + O(1/n),$$

where $\gamma \approx 0.5772156649$ is the Euler–Mascheroni constant

Coupon collector's problem - Wikipedia

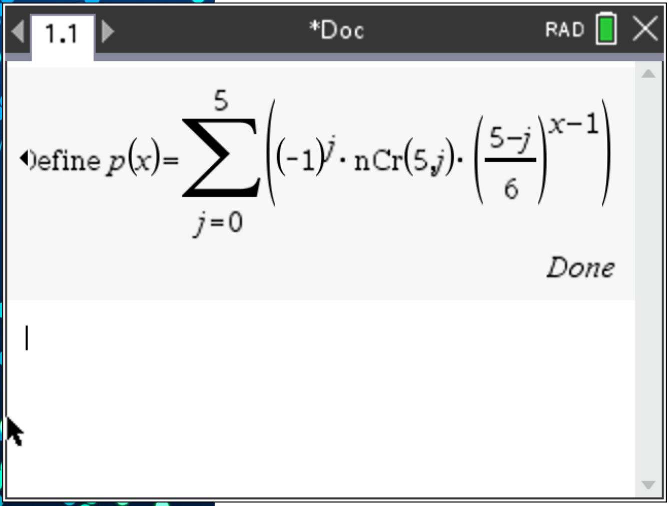
Investigation

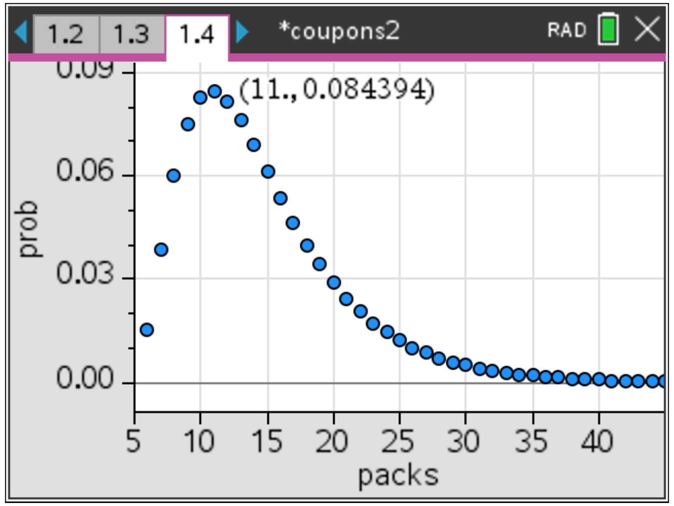
- What does the distribution look like?

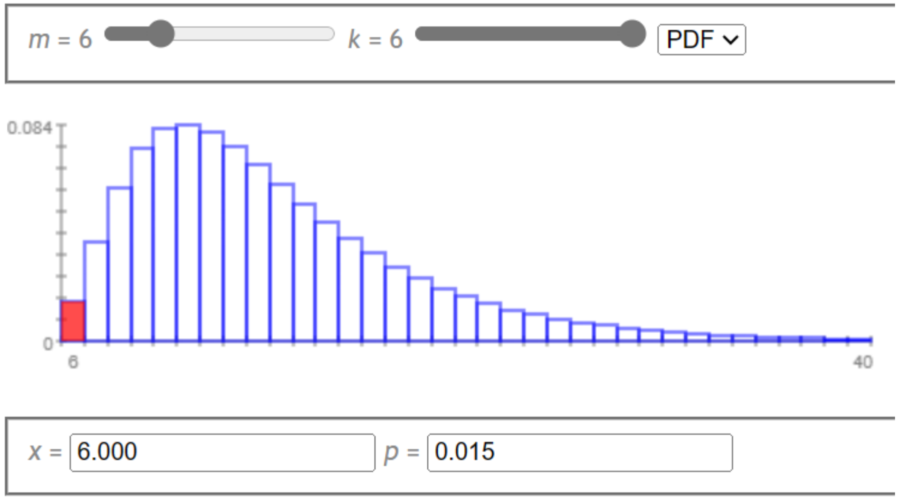
For $k \in \{1, 2, \ldots, m\}$, the probability density function of W_k is given by

$$\mathbb{P}(W_k = n) = \binom{m-1}{k-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \left(\frac{k-j-1}{m}\right)^{n-1}, \quad n \in \{k, k+1, \ldots\}$$

https://www.randomservices.org/random/urn/Coupon.html







http://www.distributome.org/V3/calc/CouponCollectorCalculator.html

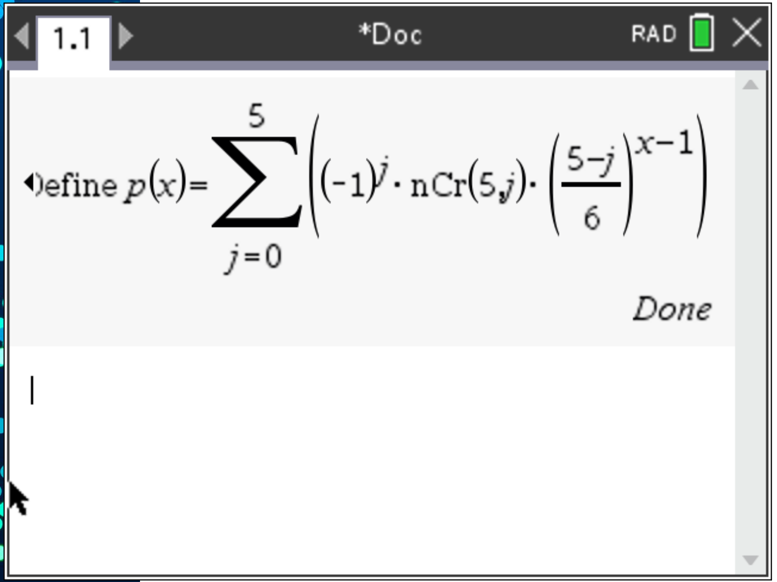
Investigation

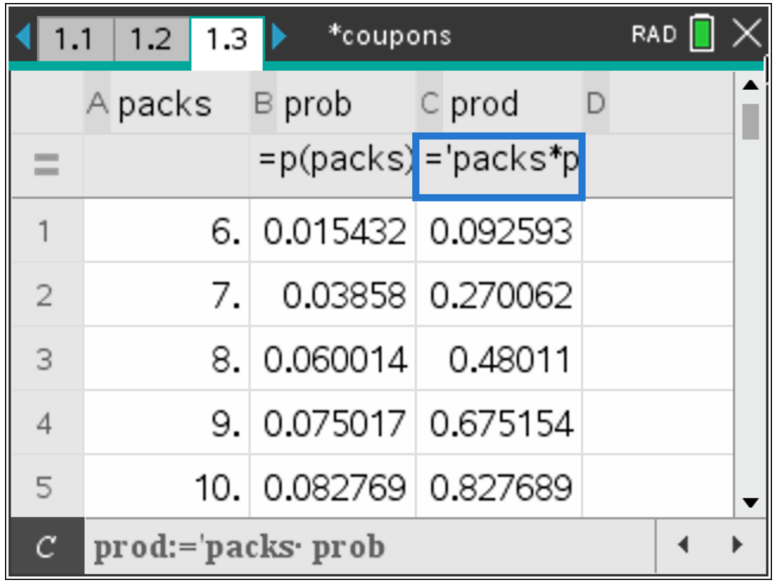
- What does the distribution look like?

For $k \in \{1, 2, \ldots, m\}$, the probability density function of W_k is given by

$$\mathbb{P}(W_k = n) = \binom{m-1}{k-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} \left(\frac{k-j-1}{m} \right)^{n-1}, \quad n \in \{k, k+1, \ldots\}$$

https://www.randomservices.org/random/urn/Coupon.html

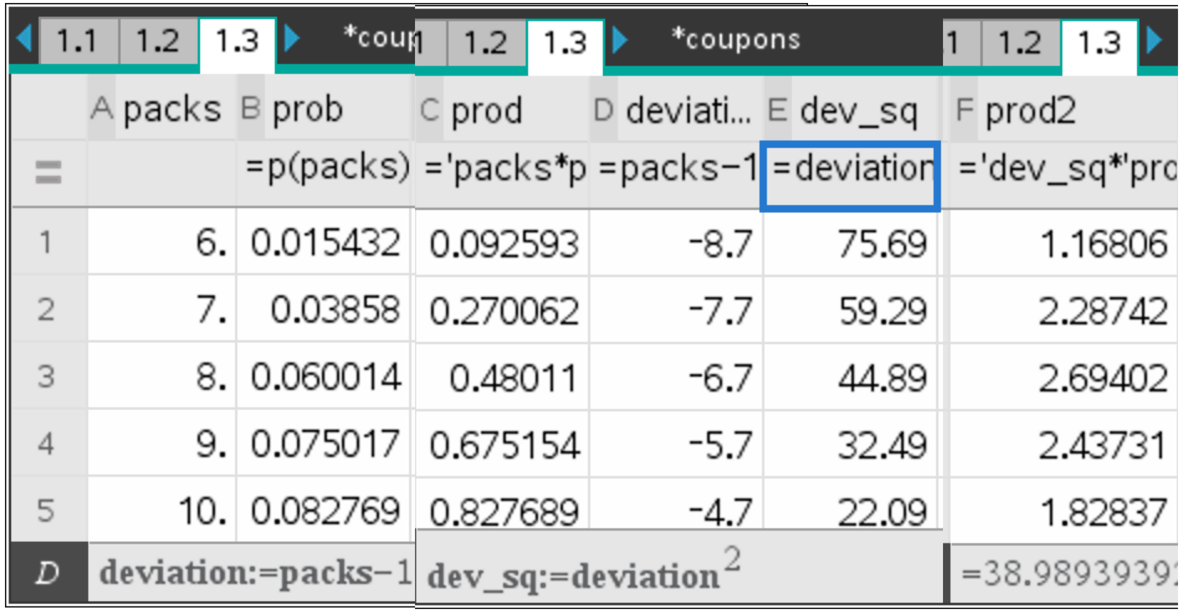


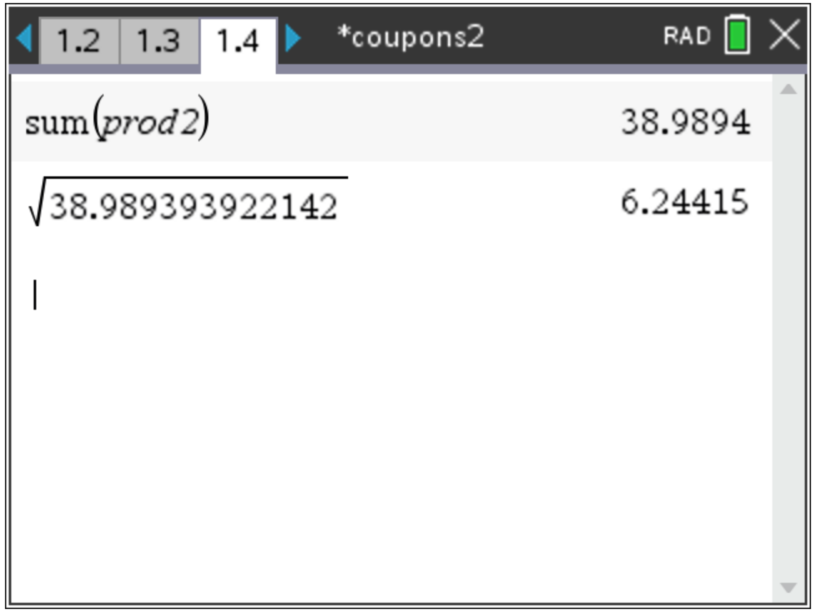


Investigation

- What does the distribution look like?

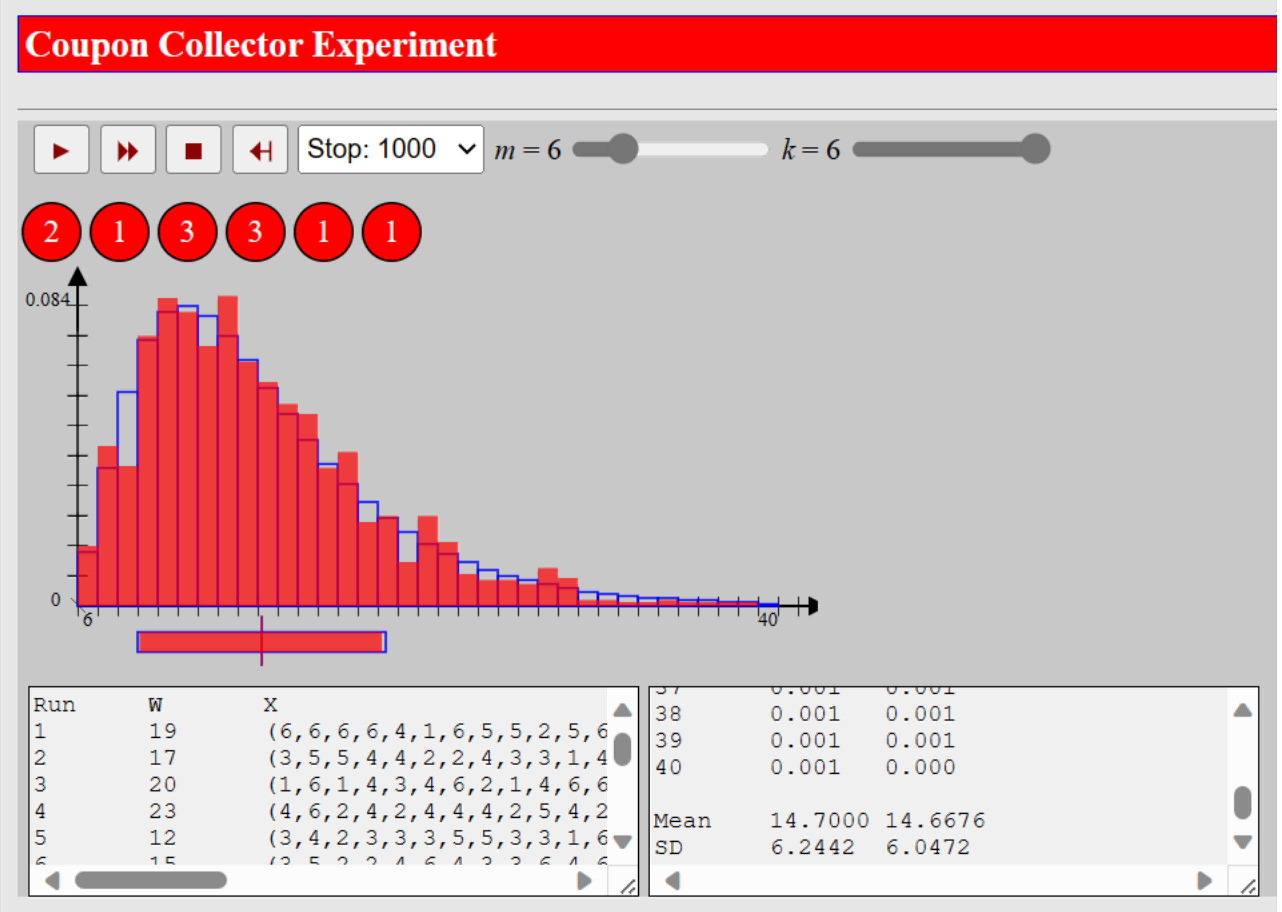
Variance
$\sigma^2 = \sum (x - \mu)^2\, p(x)$





Investigation

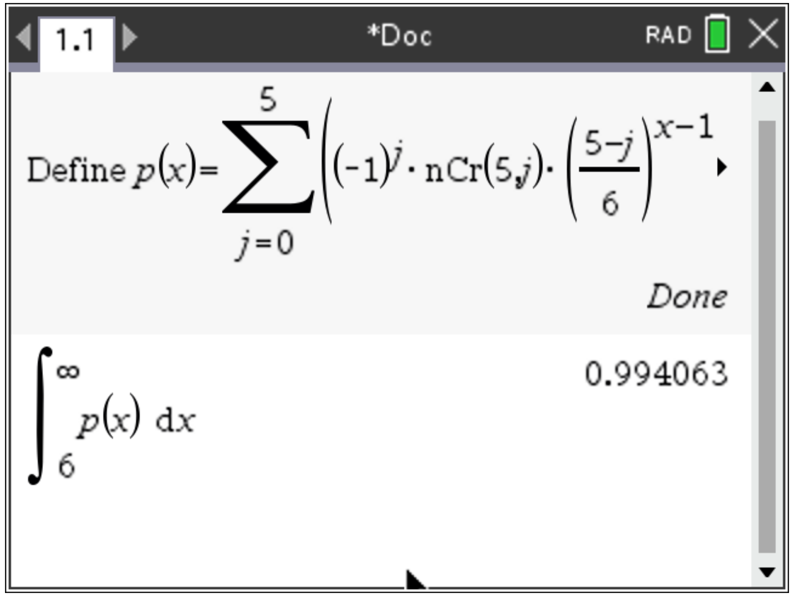
- What is the standard deviation of the number of boxes?

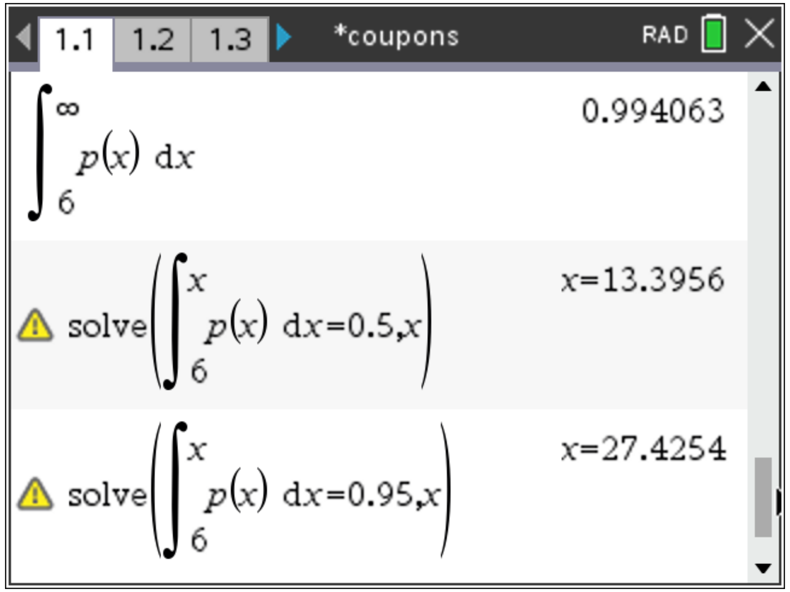


https://www.randomservices.org/random/apps/CouponCollector.html

Investigation

- We can also consider confidence levels:
  e.g. How many boxes of cereal would I need to be
  a) 50%
  b) 95%
  confident of having now collected the whole set?





Event App

App Download Instructions

Step 1: Download the App 'Arinex One' from the App Store or Google Play

App Store

Google Play

Step 2: Enter Event Code: **mav**

Step 3: Enter the email you registered with

Step 4: Enter the Passcode you receive via email and click 'Verify'. Please be sure to check your Junk Mail for the email, or see the Registration Desk if you require further assistance.

Be in it to WIN!