1. Should you sample with or without replacement? ^{D,S}

Urn A contains 2 red balls and 1 green ball, and urn B contains 101 red balls and 100 green balls. An urn is chosen at random and two balls are randomly drawn from the selected urn. You win if you correctly identify whether the selected urn was A or B.

Compute the probabilities of winning for each of the following rules and decide which gives you the highest probability of winning.

Question 1: The first ball is replaced before the second drawing.

Question 2: The first ball is not replaced before the second drawing.

Question 3: After the first ball is drawn you can decide whether it will be replaced or not.

Hint: When computing probabilities:

$$\frac{101}{201} \approx \frac{100}{201} \approx \frac{100}{200} \approx \frac{1}{2}\,.$$

Solution

There are four outcomes which we denote by RR, RG, GR, GG. For each rule compute the conditional probabilities of the four outcomes given that urn A or urn B was selected initially. These probabilities are multiplied by $1/2$ to take into account the random selection of a urn.

Answer 1: Drawing with replacement:

$$\begin{array}{lcccc}
P(RR|A) & = & \frac{2}{3}\cdot\frac{2}{3} & = & \frac{4}{9} \\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(RG|A) & = & \frac{2}{3}\cdot\frac{1}{3} & = & \frac{2}{9} \\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GR|A) & = & \frac{1}{3}\cdot\frac{2}{3} & = & \frac{2}{9} \\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GG|A) & = & \frac{1}{3}\cdot\frac{1}{3} & = & \frac{1}{9} \\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

If the outcome is RR there is a higher probability that urn A was selected $(4/9)$ than that urn B was selected $(1/4)$; otherwise, there is a higher probability that urn B was selected. Therefore:

$$P(\text{winning}) = \frac{1}{2}\left(\frac{4}{9}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}\right) = \frac{43}{72} \approx 0.5972\,.$$

Answer 2: Drawing without replacement:

$$\begin{array}{lcccc}
P(RR|A) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(RG|A) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GR|A) & = & \frac{1}{3}\cdot 1 & = & \frac{1}{3} \\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GG|A) & = & \frac{1}{3}\cdot 0 & = & 0 \\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

If the outcome is GG there is (of course!) a higher probability that urn B was selected than that urn A was selected; otherwise, there is a higher probability that urn A was selected. Therefore:

$$P(\text{winning}) = \frac{1}{2}\left(\frac{1}{3}+\frac{1}{3}+\frac{1}{3}+\frac{1}{4}\right) = \frac{5}{8} = 0.6250\,,$$

which is greater than the probability of winning when sampling with replacement.

Answer 3: The decision is based on the outcome of the first draw.

If the first drawing is from urn A the probabilities must be conditioned on the decision to sample with or without replacement. Drawing first from urn B does not affect the probabilities because of the approximation in the hint.

$$\begin{array}{lcccc}
P(RR|A,w) & = & \frac{2}{3}\cdot\frac{2}{3} & = & \frac{4}{9} \\
P(RR|A,w/o) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(RG|A,w) & = & \frac{2}{3}\cdot\frac{1}{3} & = & \frac{2}{9} \\
P(RG|A,w/o) & = & \frac{2}{3}\cdot\frac{1}{2} & = & \frac{1}{3} \\
P(RG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GR|A,w) & = & \frac{1}{3}\cdot\frac{2}{3} & = & \frac{2}{9} \\
P(GR|A,w/o) & = & \frac{1}{3}\cdot 1 & = & \frac{1}{3} \\
P(GR|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4} \\
\hline
P(GG|A,w) & = & \frac{1}{3}\cdot\frac{1}{3} & = & \frac{1}{9} \\
P(GG|A,w/o) & = & \frac{1}{3}\cdot 0 & = & 0 \\
P(GG|B) & = & \frac{1}{2}\cdot\frac{1}{2} & = & \frac{1}{4}
\end{array}$$

If a red ball is drawn first then $\frac{4}{9} > \frac{1}{4}$ and $\frac{2}{9} < \frac{1}{4}$ with replacement, whereas $\frac{1}{3} > \frac{1}{4}$ and $\frac{1}{3} > \frac{1}{4}$ without replacement, so the second ball can help identify the urn only if the drawing is done with replacement: urn A if red, urn B if green. Choose the draw with replacement and:

$$P(\text{winning if red first}) = \frac{1}{2}\left(\frac{4}{9} + \frac{1}{4}\right) = \frac{25}{72}.$$

If a green ball is drawn first then $\frac{2}{9} < \frac{1}{4}$ and $\frac{1}{9} < \frac{1}{4}$ with replacement, whereas $\frac{1}{3} > \frac{1}{4}$ and $0 < \frac{1}{4}$ without replacement, so the second ball can help identify the urn only if the drawing is done without replacement: urn A if red, urn B if green. Choose the draw without replacement and:

$$P(\text{winning if green first}) = \frac{1}{2}\left(\frac{1}{3} + \frac{1}{4}\right) = \frac{7}{24}.$$

The probability of winning is:

$$P(\text{winning}) = \frac{25}{72} + \frac{7}{24} = \frac{23}{36} \approx 0.6389.$$

The highest probability of winning is obtained when the decision to draw with or without replacement depends on the result of the first draw.

Simulation

```
With replacement:
Expectation of winning = 0.5972
Average wins           = 0.5976
Without replacement:
Expectation of winning = 0.6250
Average wins           = 0.6207
Decide after first draw:
Expectation of winning = 0.6389
Average wins           = 0.6379
```

2. The ballot box ^{S}

In an election there are two candidates A and B. A receives a votes and B receives b votes, $a > b$. The votes are counted one-by-one and the running totals $(a_i, b_i), 1 \leq i \leq a+b$ are updated as each vote is counted. What is the probability that for at least one i, $a_i = b_i$?

Question 1: Solve for $a = 3, b = 2$ by listing (a_i, b_i) for $1 \leq i \leq 5$.

Question 2: Solve for all $a > b$.

Hint 1: What can you say about the identity of the candidate who leads until the first tie occurs?

Hint 2: What is the significance of the first vote counted?

Solution

Answer 1: The number of arrangements of running totals is $\binom{5}{2} = \binom{5}{3} = 10$, because the positions of the votes for one candidate determine the positions of the votes for the other candidate. The following table lists the possible arrangements of the votes and of running totals with first ties emphasized:

A	A	A	B	B	$(1,0)$	$(2,0)$	$(3,0)$	$(3,1)$	$(3,2)$
A	A	B	A	B	$(1,0)$	$(2,0)$	$(2,1)$	$(3,1)$	$(3,2)$
A	B	A	A	B	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(3,1)$	$(3,2)$
B	A	A	A	B	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(3,1)$	$(3,2)$
A	A	B	B	A	$(1,0)$	$(2,0)$	$(2,1)$	$(\mathbf{2},\mathbf{2})$	$(3,2)$
A	B	A	B	A	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(2,2)$	$(3,2)$
B	A	A	B	A	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(2,1)$	$(2,2)$	$(3,2)$
A	B	B	A	A	$(1,0)$	$(\mathbf{1},\mathbf{1})$	$(1,2)$	$(2,2)$	$(3,2)$
B	A	B	A	A	$(0,1)$	$(\mathbf{1},\mathbf{1})$	$(1,2)$	$(2,2)$	$(3,2)$
B	B	A	A	A	$(0,1)$	$(0,2)$	$(1,2)$	$(\mathbf{2},\mathbf{2})$	$(3,2)$

There are ties in all the arrangements except for the first two so:

$$P(\text{tie occurs with } (3,2) \text{ votes}) = \frac{8}{10} = \frac{4}{5}\,.$$

Answer 2: We begin the solution with a discussion on how to approach the second question. Here are some arrangements for $(a, b) = (3, 2)$ votes until the first tie occurs:

A leads until tie	B leads until tie
A B	B A
A A B B	B B A A

For every arrangement where A leads until the first tie, there is a mirror image arrangement where B leads until the first tie. The mirror image is obtained by exchanging all A's and B's.

Before the first tie one of the candidates must be leading. If the first vote counted is for B there must be a tie later in the count since $a > b$. The probability that the first vote is for B is:

$$P(\text{first vote for } B) = \frac{b}{a+b}\,.$$

By mirroring the positions of the votes, the number of sequences resulting in a tie that begin with a vote for A is the same as the number of sequences resulting in a tie that begin with a vote for B. Therefore:

$$P(\text{tie occurs}) = 2 \cdot \frac{b}{a+b}\,.$$

Check:

$$P(\text{tie occurs with } (3,2) \text{ votes}) = 2 \cdot \frac{2}{2+3} = \frac{4}{5}\,.$$

Simulation

```
For a =  3, b =  2:
Probability of a tie = 0.8000
Proportion of ties   = 0.8118
For a = 10, b =  8:
Probability of a tie = 0.8889
Proportion of ties   = 0.8977
For a = 20, b = 18:
Probability of a tie = 0.9474
Proportion of ties   = 0.9354
```

3. Ties in matching pennies ^{D,S}

Toss a pair of fair coins N times, N even, and keep count of how many times the parity is even (heads-heads or tails-tails) and how many times the parity is odd (heads-tails or tails-heads). What is the probability of obtaining a tie (not counting the $0-0$ tie at the start)?

Question 1: Solve for $N=4$ by writing out all the possible outcomes.

Question 2: Solve for $N=6$ by developing a formula for the probability.

Question 3: Develop a formula for arbitrary even N.

Question 4: Explain why the probability for the odd number $N+1$ is the same as the probability for the even number N.

Hint: Use the solution of Problem 22.

Solution

Answer 1: Denote tosses with even parity by E and tosses with odd parity by O. Ten out of the sixteen arrangements of tosses have ties (emphasized):

EEEE	EEEO	EEOE	EEOO	EOEE	EOEO	EOOE	EOOO
OEEE	OEEO	OEOE	OEOO	OOEE	OOEO	OOOE	OOOO

Answer 2: By Problem 22:

$$P(\text{tie on toss } i) = \begin{cases} 2i/N & \text{if } i \leq N/2 \\ 2(N-i)/N & \text{if } i \geq N/2, \end{cases} \tag{1}$$

since the ballot box problem showed that the smaller value determines the probability.

The following computations are quite complex so we justify each step in detail.

The probability of i evens is given by the binomial distribution:

$$P(i \text{ evens}) = \binom{N}{i}\left(\frac{1}{2}\right)^{i}\left(\frac{1}{2}\right)^{N-i} = \binom{N}{i}\left(\frac{1}{2}\right)^{N} = 2^{-N}\binom{N}{i}. \tag{2}$$

The probability of a tie is the sum over i of the probability of obtaining i evens times the probability of a tie on the ith toss (Equation ??). For $N = 6$:

$$P(\mathsf{ties}) = 2 \cdot 2^{-6} \left[\frac{0}{6}\binom{6}{0} + \frac{1}{6}\binom{6}{1} + \frac{2}{6}\binom{6}{2} + \frac{3}{6}\binom{6}{3} + \frac{2}{6}\binom{6}{4} + \frac{1}{6}\binom{6}{5} + \frac{0}{6}\binom{6}{6} \right] . \quad (3)$$

Equation ?? follows from Equation ?? by deleting the two zero terms, expressing the combinations as factorials, canceling $1/6$ from $6!$:

$$P(\mathsf{ties}) = 2^{-5} \left[1 \cdot \frac{5!}{1!5!} + 2 \cdot \frac{5!}{2!4!} + 3 \cdot \frac{5!}{3!3!} + 2 \cdot \frac{5!}{4!2!} + 1 \cdot \frac{5!}{5!1!} \right] . \quad (4)$$

Equation ?? is obtained by canceling i from $i!$:

$$P(\mathsf{ties}) = 2^{-5} \left[\frac{5!}{1!5!} + \frac{5!}{1!4!} + \frac{5!}{2!3!} + \frac{5!}{4!1!} + \frac{5!}{5!1!} \right] . \quad (5)$$

To obtain Equation ?? from Equation ?? add and subtract $\frac{5!}{3!2!}$:

$$P(\mathsf{ties}) = 2^{-5} \left[\left(\frac{5!}{1!5!} + \frac{5!}{1!4!} + \frac{5!}{2!3!} + \frac{5!}{3!2!} + \frac{5!}{4!1!} + \frac{5!}{5!1!} \right) - \frac{5!}{3!2!} \right] . \quad (6)$$

Equation ?? results from replacing $1!$ by $0!$:

$$P(\mathsf{ties}) = 2^{-5} \left[\left(\frac{5!}{0!5!} + \frac{5!}{1!4!} + \frac{5!}{2!3!} + \frac{5!}{3!2!} + \frac{5!}{4!1!} + \frac{5!}{5!0!} \right) - \frac{5!}{3!2!} \right] . \quad (7)$$

By expressing the factorials back as combinations we obtain Equation ??:

$$P(\mathsf{ties}) = 2^{-5} \left[\binom{5}{0} + \binom{5}{1} + \binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5} - \binom{5}{3} \right] . \quad (8)$$

Finally, Equation ?? results from the binomial theorem:

$$P(\mathsf{ties}) = 2^{-5} \left(2^5 - 10 \right) = \frac{11}{16} \approx 0.6875 . \quad (9)$$

Answer 3: Perform the same calculations as in Answer 2: but using arbitrary N. The result is:

$$P(\mathsf{ties}) = 2^{-N+1} \left[2^{N-1} - \binom{N-1}{N/2} \right] = \left[1 - \binom{N-1}{N/2} \Big/ 2^{N-1} \right] .$$

Answer 4: The first tie on the $N+1$'st toss occurs only if the counts are nearly equal after the Nth toss:

$$((N/2) - 1, (N/2) + 1)$$
$$((N/2), (N/2))$$
$$((N/2) + 1, (N/2) - 1)$$

but whatever the outcome of the final toss the counts will not be equal.

Simulation

```
For  4 tosses:
Probability of ties = 0.6250
Proportion of ties  = 0.6192
For  6 tosses:
Probability of ties = 0.6875
Proportion of ties  = 0.6900
For  7 tosses:
Probability of ties = 0.6875
Proportion of ties  = 0.6811
For 10 tosses:
Probability of ties = 0.7539
Proportion of ties  = 0.7559
For 20 tosses:
Probability of ties = 0.8238
Proportion of ties  = 0.8255
```

5. Lengths of random chords ^{S}

Select a random chord in the unit circle. What is the probability that the length of the chord is greater than 1?

To solve the problem you first have to decide what "select a random chord" means. Solve the problem for each of the following possibilities:

Question 1: The distance of the chord from the center is uniformly distributed.

Question 2: The midpoint of the chord is uniformly distributed within the circle.

Question 3: The endpoints of the chord are uniformly distributed on the circumference of the circle.

Solution

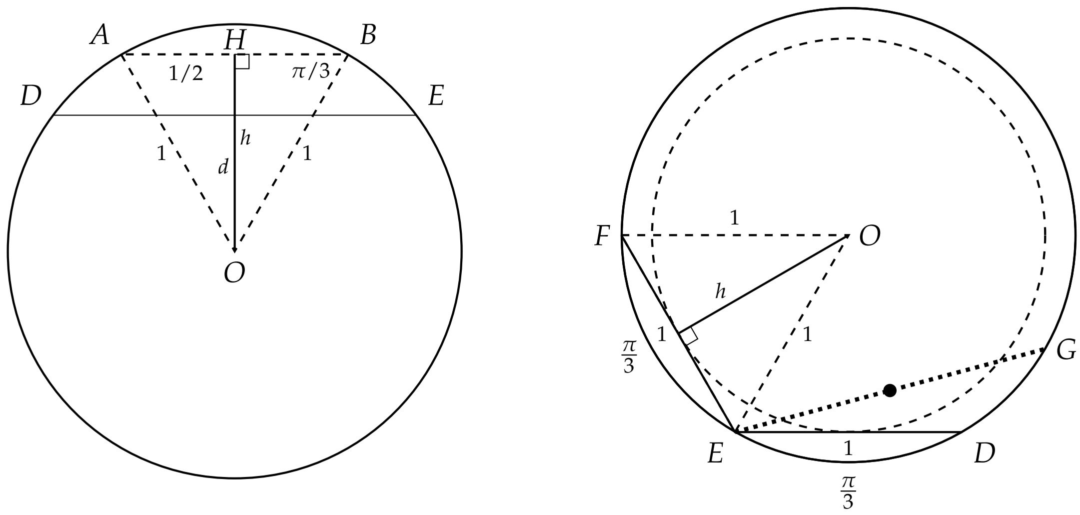
Answer 1: A chord is larger than the radius if it is closer to the center than a chord of length 1. Let $\overline{AB}$ be a chord of length 1 and construct the altitude $\overline{OH}$ from the center O to the chord (Figure ??). Since $\triangle AOB$ is equilateral, $\triangle OHB$ is a right triangle and the length of the altitude is:

$$h = \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}\,.$$

Let d be the distance of a chord $\overline{DE}$ from the center. By assumption d is uniformly distributed in $(0, 1)$ so:

$$P(\overline{DE} > 1) = P(d < h) = \frac{h}{1} = \frac{\sqrt{3}}{2} \approx 0.866\,.$$

Answer 2: Construct a circle with center O and radius h, where h is the length of the altitude to a chord of length 1. A tangent to any point on this circle will be a chord $\overline{FE}$ whose length



(a) Distance of chord from center uniformly distributed in $(0, 1)$

(b) Midpoint of chord uniformly distributed within the circle and endpoints of chord uniformly distributed on the circumference

is 1. Any chord $\overline{EG}$ whose midpoint is within this circle will have a length greater than 1 (Figure ??). The probability that the length of the chord is greater than 1 is therefore the ratio of the areas of the two circles:

$$P(\overline{EG} > 1) = \frac{\pi \cdot h^2}{\pi \cdot 1^2} = h^2 = \frac{3}{4}.$$

This is the square of the probability computed in the previous question.

Answer 3: Choose an arbitrary point on the circumference of the unit circle (E in Figure ??). Any other point on the circumference (such as G in the Figure) determines a chord whose length is greater than one unless that point falls on the arcs $\widehat{EF}$ or $\widehat{ED}$. The probability is therefore the ratio of the arc $\widehat{FD}$ to the circumference of the unit circle:

$$P(\overline{EG} > 1) = \frac{(2\pi - (2\pi/3)) \cdot 1}{2\pi \cdot 1} = \frac{2}{3}.$$

Simulation

The simulation is for choosing two random points on the circumference.

```
Probability of long chords = 0.6667
Proportion of long chords  = 0.6627
```

6. The hurried duelers ^{S}

A and B arrive at a meeting point at a random time with uniform distribution within a one-hour period. If A arrives first and B does not arrive within 5 minutes, A leaves. Similarly if B arrives first and A does not arrive within 5 minutes, B leaves. What is the probability that they meet?

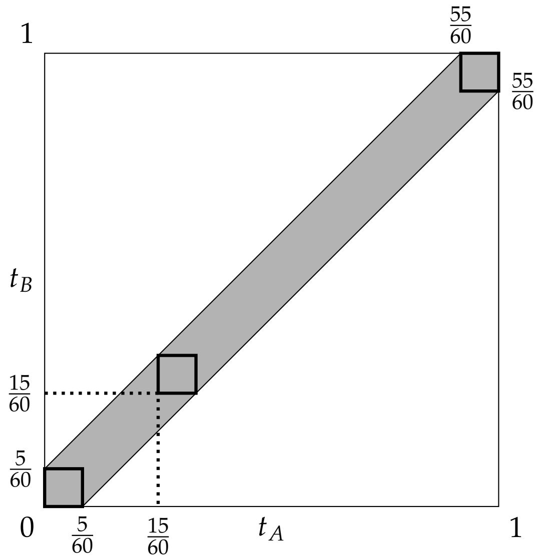


Figure 2: Times that ensure a meeting between A and B

Time within the one-hour period is continuous in the range $[0,1]$, that is, you cannot count a discrete number of minutes or seconds to compute probabilities. You can compute the probabilities of durations.

Hint: Draw a graph with A's time of arrival as the the x-axis and B's time of arrival as the y-axis.

Solution

Without loss of generality assume that A arrives first. If A arrives at $t_A = 0$ and if B arrives before $t_B = 5/60$ they meet, otherwise they do not. This is shown in Figure ?? by the small square at the origin. If A arrives later then B also has to arrive later by the same amount; for example, if A arrives at $t_A = 15$, B must arrive between $t_B = 15$ and $t_B = 20$. Therefore, the meeting will take place during a square of time obtained by moving the square by 15 from $(0,0)$ to $(15/60, 15/60)$.

The probability of a meeting is the ratio of the area of the graph colored gray to the area of the large square. It is easier to compute the complement which is the ratio of the area of the two white triangles to the area of the large square:

$$\begin{aligned} P(A, B \text{ meet}) &= 1 - P(A, B \text{ don't meet}) \\ &= 1 - 2 \cdot \left(\frac{1}{2} \cdot \frac{55}{60} \cdot \frac{55}{60}\right) = \frac{23}{144} \approx 0.1597\,. \end{aligned}$$

Simulation

```
Probability of meeting    = 0.1597
Proportion of meetings    = 0.1549
```

7. Catching the cautious counterfeiter ^{S}

There are n boxes each with n coins one of which is counterfeit. Draw one coin from each box and test it to determine whether it is counterfeit or genuine. What is the probability that all the coins that are drawn are are real?

Question 1: Solve for $n = 10$.

Question 2: Solve for $n = 100$.

Question 3: Solve for arbitrary n.

Question 4: Develop a formula for limit of the probability as n tends to infinity.

Solution

The draws are independent so the probability is the product of the probabilities for each draw.

Answer 1:

$$P(\textsf{all coins are genuine}) = \left(\frac{9}{10}\right)^{10} = 0.3487\,.$$

Answer 2:

$$P(\textsf{all coins are genuine}) = \left(\frac{99}{100}\right)^{100} = 0.3660\,.$$

Answer 3:

$$P(\textsf{all coins are genuine}) = \left(\frac{n-1}{n}\right)^{n}\,.$$

Answer 4:

$$\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \frac{1}{e} \approx 0.3679\,. \tag{10}$$

This limit can be proved using differential calculus. First we compute the limit of the natural logarithm of the lefthand side of Equation ??:

$$\lim_{n\to\infty} \ln\left(1-\frac{1}{n}\right)^n = \lim_{n\to\infty} n\ln\left(1-\frac{1}{n}\right) = \lim_{n\to\infty}\frac{\ln\left(1-\frac{1}{n}\right)}{1/n}\,.$$

Taking the limit gives $(ln\ 1)/0 = 0/0$ but by l'Hôpital's rule we can replace expression by the quotient of the derivatives:

$$\begin{aligned}\lim_{n\to\infty} \ln\left(1-\frac{1}{n}\right)^n &= \lim_{n\to\infty}\frac{\left(1-\frac{1}{n}\right)^{-1}\left(-(-n^{-2})\right)}{-n^{-2}} = -1\\ \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n &= e^{-1}\,.\end{aligned}$$

Simulation

```
For  10 boxes:
Probability of all real = 0.3487
Proportion all real     = 0.3480
For 100 boxes:
Probability of all real = 0.3660
Proportion all real     = 0.3730
For 200 boxes:
Probability of all real = 0.3670
Proportion all real     = 0.3690
```

8. Catching the greedy counterfeiter ^{S}

There are n boxes each with n coins m of which are counterfeit. Draw one coin from each box and test it to determine whether it is counterfeit or not. What is the probability $P(n,m,r)$ that r of the coins that are drawn are counterfeit?

Question 1: Develop a formula for $P(n,m,r)$.

Question 2: Compute $P(20,10,2), P(20,10,8), P(20,5,2), P(20,5,4)$.

Solution

Answer 1: There are $\binom{n}{r}$ choices of boxes from which the counterfeit coins can be drawn. From the binomial distribution:

$$P(n,m,r) = \binom{n}{r}\left(\frac{m}{n}\right)^r\left(\frac{n-m}{n}\right)^{n-r}.$$

Answer 2:

$$\begin{aligned} P(20,10,2) &= \binom{20}{2}\left(\frac{10}{20}\right)^2\left(\frac{10}{20}\right)^{18} \approx 0.0002 \\ P(20,10,8) &= \binom{20}{8}\left(\frac{10}{20}\right)^8\left(\frac{10}{20}\right)^{12} \approx 0.1201 \\ P(20,5,2) &= \binom{20}{2}\left(\frac{5}{20}\right)^2\left(\frac{15}{20}\right)^{18} \approx 0.0669 \\ P(20,5,4) &= \binom{20}{4}\left(\frac{5}{20}\right)^4\left(\frac{15}{20}\right)^{12} \approx 0.1952\,. \end{aligned}$$

Mosteller shows that for given m, r, as n tends to infinity:

$$\lim_{n\to\infty} P(n,m,r) = \frac{e^{-m}m^r}{r!}\,. \tag{11}$$

Simulation

```
For 10 bad coins,  2 draws:
Probability of counterfeit  = 0.0002
Proportion counterfeit      = 0.0002
For 10 bad coins,  8 draws:
Probability of counterfeit  = 0.1201
Proportion counterfeit      = 0.1181
For  5 bad coins,  2 draws:
Probability of counterfeit  = 0.0669
Proportion counterfeit      = 0.0688
For  5 bad coins,  4 draws:
Probability of counterfeit  = 0.1897
Proportion counterfeit      = 0.1905
```

9. Moldy gelatin ^{S}

A rectangular plate is divided into n small squares. There are an average of r microbes in each square.

Question 1: Develop a formula for probability that there are exactly r microbes in the n squares.

Question 2: Compute the probability for $n = 100, r = 3$.

Hint: This problem is similar the Problem 28.

Solution

Answer 1: Let p be the probability that a single square contains a microbe. (Ignore the possibility that a microbe is partially contained within two or more squares.) m, the average number of microbes per square, is the number of squares n times the probability p that a square contains a microbe. $P(n, m, r)$, the probability that there are exactly r microbes in the n squares is given by the binomial distribution:

$$P(n, m, r) = \binom{n}{r} \left(\frac{m}{n}\right)^{r} \left(\frac{n-m}{n}\right)^{n-r} .$$

Answer 2:

$$P(10, 3, 3) = \binom{100}{3} \left(\frac{3}{100}\right)^{3} \left(\frac{97}{100}\right)^{97} \approx 0.2275 .$$

Equation ?? also applies here:

$$\lim_{n\to\infty} P(n, 3, 3) = \frac{e^{-3} \cdot 3^{3}}{3!} \approx 0.2240 .$$

Simulation

```
For  20 squares:
Probability of exactly  3 microbes  = 0.2428
Proportion of exactly   3 microbes  = 0.2436
Probability of exactly  5 microbes  = 0.2023
Proportion of exactly   5 microbes  = 0.1954
For 100 squares:
Probability of exactly  3 microbes  = 0.2275
Proportion of exactly   3 microbes  = 0.2247
Probability of exactly  5 microbes  = 0.1800
Proportion of exactly   5 microbes  = 0.1851
```