11/3/14

Exam 2 : Total = 40
Average = 32.3
Quartiles = 27, 35, 38
St. Deviation = 7.

Exam 3 is on Mon Dec 8.

New Topic : Probabity

Recall : The Binomial Theorem says that for all numbers a & b and integers $n \geq 0$ we have

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Now let p & q be numbers such that

- $0 \leq p \leq 1$
- $0 \leq q \leq 1$
- $p + q = 1$.

The Binomial Theorem says

$$1 = 1^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}.$$

What does this mean?

Suppose we have a <u>biased</u> <u>coin</u> with

$$\text{Prob}(\text{heads}) = p$$
$$\text{Prob}(\text{tails}) = q.$$

If we flip the coin n times, what is the probability we will get heads exactly k times?

Example: If we flip the coin 3 times, the probability of getting the sequence HTH is

$$\begin{aligned} \text{Prob}(HTH) &= \text{Prob}(H)\cdot\text{Prob}(T)\cdot\text{Prob}(H) \\ &= pqp \\ &= p^2q. \end{aligned}$$

To find the probability of "2 heads" we have to sum the different ways it can happen.

$$\text{Prob}(\text{"2 heads in 3 tosses"})$$

$$= \text{Prob}(\{HHT, HTH, THH\})$$

$$= \text{Prob}(HHT) + \text{Prob}(HTH) + \text{Prob}(THH)$$

$$= ppq + pqp + qpp$$

$$= 3p^2q.$$

In general, the probability of getting exactly k heads in n tosses of the coin is

$$\binom{n}{k} p^k q^{n-k}.$$

Q: What is the probability of getting <u>some</u> number of heads?

A: We sum over all possible k:

$$\text{Prob}(\text{"some number of heads in } n \text{ tosses"})$$

$$= \sum_k \text{Prob}(\text{"}k\text{ heads in } n \text{ tosses"})$$

$$= \sum_k \binom{n}{k} p^k q^{n-k}$$

$$= (p+q)^n = 1^n = 1,$$

as we expect. ///

Suppose in a certain population each birth has

$$\text{Prob(boy)} = 1/3 \qquad \left(\frac{1}{3} + \frac{2}{3} = 1 \right)$$
$$\text{Prob(girl)} = 2/3$$

If a certain family has 4 children, how many boys are they likely to have?

We assume this is just like a coin flip, so the probability of k boys is

$$\binom{4}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{4-k}.$$

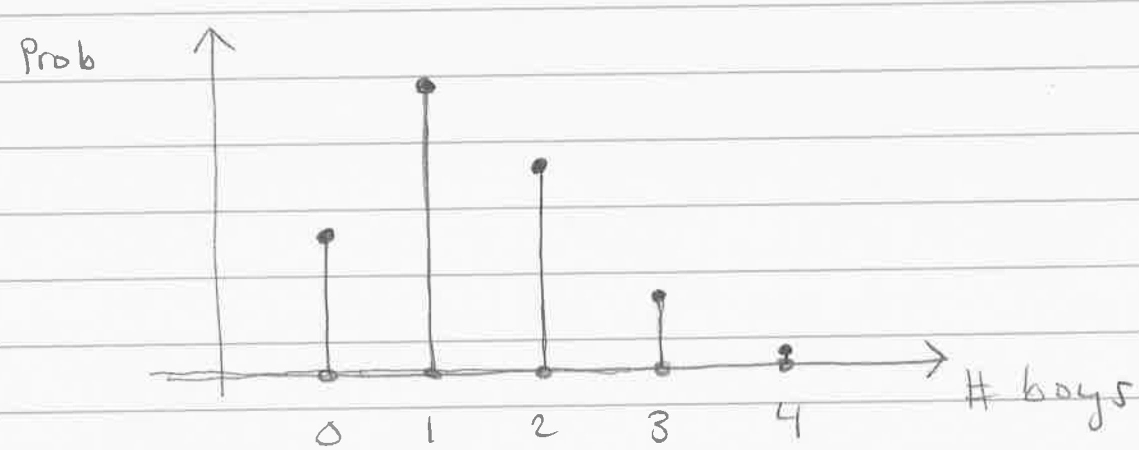
Let's compute the full distribution:

k	0	1	2	3	4
Prob(k boys)	$\binom{4}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^4$	$\binom{4}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^3$	$\binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2$	$\binom{4}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^1$	$\binom{4}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^0$
	$\frac{1\cdot 1^0\cdot 2^4}{3^4}$	$\frac{4\cdot 1^1\cdot 2^3}{3^4}$	$\frac{6\cdot 1^2\cdot 2^2}{3^4}$	$\frac{4\cdot 1^3\cdot 2^1}{3^4}$	$\frac{1\cdot 1^4\cdot 2^0}{3^4}$
	$\frac{16}{81}$	$\frac{32}{81}$	$\frac{24}{81}$	$\frac{8}{81}$	$\frac{1}{81}$

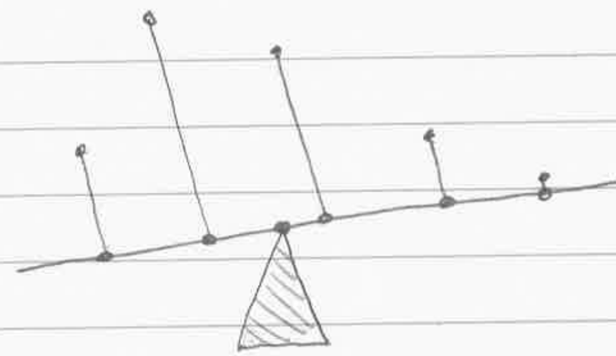
Notice that

$$\frac{16}{81}+\frac{32}{81}+\frac{24}{81}+\frac{8}{81}+\frac{1}{81}=\frac{81}{81}=1,$$

as it should be. (The probability of <u>something</u> happening is 1.) We can think of probability as a distribution of "mass":



The "average" or "expected value" is the same as the "center of mass":



Where does it balance? Archimedes tells us the answer. His "law of the lever" says that

mass x (distance from center)

should be in balance. For example, suppose we have

The system will balance when

$$m_1 d_1 = m_2 d_2 .$$

More generally if we have masses $m_1, \dots, m_n$ at positions $x_1, \dots, x_n$, then the center of mass $\overline{x}$ satisfies

$$\sum_{k=1}^{n} m_k (x_k - \overline{x}) = 0$$

$$\sum_k m_k x_k - \overline{x} \sum m_k = 0$$

$$\overline{x} = \frac{\sum_k m_k x_k}{\sum_k m_k} .$$

In our case $m_k = \text{Prob}(k \text{ boys})$ and $x_k = k$. We also have $\sum m_k = 1$. So the expected number of boys is

$$\overline{k} = \sum \text{Prob}(k \text{ boys}) \cdot k$$

$$= 0 \cdot \frac{16}{81} + 1 \cdot \frac{32}{81} + 2 \cdot \frac{24}{81} + 3 \cdot \frac{8}{81} + 4 \cdot \frac{1}{81}$$

$$= \frac{0 + 32 + 48 + 24 + 4}{81} = \frac{108}{81} = \frac{4}{3} .$$

$$= 1.33\dots$$

Does that surprise you? It shouldn't.

In a population with Prob(boy) = 1/3 and Prob(girl) = 2/3 we expect that 1/3 of all children will be boys. So in a family of 4 children we expect

$$\frac{1}{3} \cdot 4 = 1.33\ldots \quad \text{boys.}$$

In general, we have

☆ Theorem: Consider a biased coin with $P(\text{heads}) = p$, $P(\text{tails}) = 1-p$. If we toss the coin n times the expected number of heads is

$$pn.$$

Proof: By Archimedes' principle, the expected number of heads is

$$\bar{k} = \sum_{k=0}^{n} k \cdot \text{Prob}(k \text{ heads})$$

$$= \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}.$$

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

Does this simplify? Not immediately. First we must use the fact that

$$k \binom{n}{k} = n \binom{n-1}{k-1} \qquad (\text{Exam } 2.5)$$

Then we have

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

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

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

$$= pn \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j}$$

$$= pn \left(p + (1-p)\right)^{n-1}$$

$$= pn \, (1)^{n-1} = pn. \qquad ///$$

Is there an easier way to do that?
Yes, but it requires a bit more technology.

Let $B(1,p)$ be a random number that is 1 with probability p and 0 with probability $1-p$. This is called a

"Bernoulli random variable"

We can show that the sum of n Bernoulli random variables

$$B(n,p) := \underbrace{B(1,p) + B(1,p) + \cdots + B(1,p)}_{n \text{ times}}$$

takes value k with probability

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

We call $B(n,p)$ a "binomial random variable".

Now we are interested in the <u>expected</u> <u>value</u> $E(B(n,p))$.

We can use a principle called "linearity of expectation" to show that

$$E(B(n,p)) = E(B(1,p) + \cdots + B(1,p))$$

$$= E(B(1,p)) + \cdots + E(B(1,p))$$

$$= \underbrace{p + p + \cdots + p}_{n \text{ times}}$$

$$= np.$$

Here we used the fact that

$$E(B(1,p)) = p.$$

Why is this true ?

11/5/14

No HW 5 yet.
Exam 3 on Mon Dec 8.

Consider a biased coin with

$$P(\text{heads}) = p$$
$$P(\text{tails}) = 1-p.$$

If you flip the coin n times, then the probability that you get heads exactly k times is

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

But what <u>is</u> probability?

Here is a brief history.

- 1654: The Chevalier de Méré asks Pascal for help with a gambling problem. Pascal enlists the help of Fermat.

- 1812: Laplace publishes first textbook on probability theory.

- 1933 : Kolmogorov gives the first mathematical definition of probability.

Here is Kolmogorov's definition of probability. We can view it as an add-on to the theory of Boolean algebra.

Definition: A probability space is a set S (called the "sample space") together with a function

$$P : \mathcal{P}(S) \longrightarrow \mathbb{R}.$$

Subsets $E \subseteq S$ are called events. We call the real number $P(E)$ the probability of the event E. The function P must satisfy 3 rules:

1. For all events $E \subseteq S$ we have

$$0 \leq P(E) \leq 1.$$

2. $P(S) = 1$

"The probability that something happens is 1."

(3) If the events E, F are mutally exclusive (i.e., if $E \cap F = \emptyset$) then we have

$$P(E \cup F) = P(E) + P(F).$$

That's all. //

Notation: Given any subsets A, B of a set U, we say that A, B are disjoint when $A \cap B = \emptyset$. When A, B are disjoint we will use the notation

$$A \sqcup B := A \cup B,$$

and we call this the disjoint union of A and B.

Then we can rephrase the 3rd axiom by saying that

$$P(E \sqcup F) = P(E) + P(F).$$

Q: What if E & F are not disjoint?

A: Let's prove a few theorems before discussing that.

(4) For all $E \subseteq S$ we have

$$P(E^c) = 1 - P(E).$$

<u>Proof</u>: Since $S = E \sqcup E^c$, we have

$$\begin{aligned} 1 &= P(S) && (2) \\ &= P(E \sqcup E^c) \\ &= P(E) + P(E^c) && (3) \end{aligned}$$

///

(5) $P(\emptyset) = 0$

<u>Proof</u>: Since $\emptyset = S^c$ we have

$$\begin{aligned} P(\emptyset) &= P(S^c) \\ &= 1 - P(S) && (4) \\ &= 1 - 1 && (2) \\ &= 0. \end{aligned}$$

///

{

(6) If $E \subseteq F$ then $P(E) \leq P(F)$.

<u>Proof</u>: If $E \subseteq F$ then we can write F as a disjoint union $F = E \sqcup (E^c \cap F)$:

Then we have

$$\begin{aligned} 0 &\leq P(E^c \cap F) && (1) \\ P(E) &\leq P(E) + P(E^c \cap F) \\ P(E) &\leq P(E \sqcup (E^c \cap F)) \\ P(E) &\leq P(F) && (3) \end{aligned}$$

///

More Notation: If the sets $A_1, A_2, \ldots, A_n$ are pairwise disjoint (i.e. if $A_i \cap A_j = \emptyset$ for all $i \neq j$) then we will write

$$\bigsqcup_{i=1}^{n} A_i := \bigcup_{i=1}^{n} A_i$$

and we'll call this the <u>disjoint union</u>.

(7) For any events $E_1, E_2, \ldots, E_n$ that are pairwise disjoint, we have

$$P\left(\bigsqcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i).$$

<u>Proof</u>: We will exercise our induction skills.

- First, note that the statement is true when $n=2$. (This is just axiom (3).)

- Next, <u>assume</u> for induction that the statement is true for n. We will show that it remains true for $n+1$.

So consider any events $E_1, E_2, \ldots, E_{n+1}$ that are pairwise disjoint. We want to show that

$$P\left(\bigsqcup_{i=1}^{n+1} E_i\right) = \sum_{i=1}^{n+1} P(E_i).$$

To do this, we will write

$$\bigsqcup_{i=1}^{n+1} E_i = E_1 \sqcup E_2 \sqcup \cdots \sqcup E_{n+1}$$

$$= (E_1 \sqcup E_2 \sqcup \cdots \sqcup E_n) \sqcup E_{n+1}.$$

$$= \left(\bigsqcup_{i=1}^{n} E_i\right) \sqcup E_{n+1}$$

Then axiom ③ says

$$P\left(\bigsqcup_{i=1}^{n+1} E_i\right) = P\left(\left(\bigsqcup_{i=1}^{n} E_i\right) \sqcup E_{n+1}\right)$$
$$= P\left(\bigsqcup_{i=1}^{n} E_i\right) + P(E_{n+1})$$

and our induction assumption says

$$P\left(\bigsqcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i).$$

Putting these together gives

$$P\left(\bigsqcup_{i=1}^{n+1} E_i\right) = P\left(\bigsqcup_{i=1}^{n} E_i\right) + P(E_{n+1})$$
$$= \left(\sum_{i=1}^{n} P(E_i)\right) + P(E_{n+1})$$
$$= \sum_{i=1}^{n+1} P(E_i).$$

We are done by induction. ///

[Remark: This is a very typical argument in mathematics. I'll ask for a similar proof on HW 5.]

Finally, we have a very important theorem.

☆ Principle of Inclusion-Exclusion:

For *any* events $E, F \subseteq S$ (not necessarily disjoint) we have

$$\boxed{P(E \cup F) = P(E) + P(F) - P(E \cap F)}$$

Proof: Note that we have disjoint unions

(a) $E \cup F = (E^c \cap F) \sqcup (E \cap F^c) \sqcup (E \cap F)$
(b) $E = (E \cap F) \sqcup (E \cap F^c)$
(c) $F = (E \cap F) \sqcup (E^c \cap F)$

Applying (3) and (7) gives

(a) $P(E \cup F) = P(E^c \cap F) + P(E \cap F^c) + P(E \cap F)$
(b) $P(E) = P(E \cap F) + P(E \cap F^c)$
(c) $P(F) = P(E \cap F) + P(E^c \cap F)$.

Now add $P(E \cap F)$ to both sides of (a) and apply (b) and (c) to get.

$P(E \cup F) + P(E \cap F)$

$$= \underbrace{P(E \cap F^c) + P(E \cap F)} + \underbrace{P(E^c \cap F) + P(E \cap F)}$$

$$= P(E) + P(F).$$

An important special case of P.I.E. :

Let S be a finite set. Then the function

$$P : \wp(S) \longrightarrow \mathbb{R}$$
$$E \longmapsto \#E / \#S$$

satisfies Kolmogorov's axioms ①, ②, ③.
Hence this function also satisfies P.I.E. :
For all $E, F \subseteq S$ we have

$$\frac{\#(E \cup F)}{\#S} = \frac{\#E}{\#S} + \frac{\#F}{\#S} - \frac{\#(F \cap E)}{\#S}$$

Multiply both sides by $\#S$ to get

$\#(E \cup F) = \#E + \#F - \#(E \cap F)$.

Do you recognize this?

Example: Roll two fair 6-sided dice.

Suppose one is red and one is blue, so we can tell them apart. The sample space for this experiment is

$$S = \{1, 2, 3, 4, 5, 6\}^2$$

$$= \{ (1,1), (1,2), \ldots, (1,6), \\ (2,1), (2,2), \ldots, (2,6), \\ \vdots \\ (6,1), (6,2), \ldots, (6,6) \}.$$

Note that $\#S = 6^2 = 36$.

Since the dice are <u>fair</u>, each of the 36 outcomes is equally likely, and so the probability of an event $E \subseteq S$ is just

$P(E) = \#E / \#S = \#E / 36.$

What is the probability of "rolling a 6"?

We have $E =$ "rolling a 6"

$= \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$

so the probability is

$$P(E) = \#E / \#S$$

$$= \frac{5}{36}.$$

On average we will "roll a 6" in 5 out of every 36 rolls of the dice.

<u>Example</u>: Flip a biased coin 4 times. Assume that $P(H) = p$.

What is the probability that you get heads for the first time on the 3rd flip?

The sample space is

$$S = \{H, T\}^4$$

The relevant event is

$$E = \{TTHT, TTHH\}.$$

In this experiment the outcomes are <u>NOT</u> equally likely, so just counting things is not enough.

We have

$$\begin{aligned} P(E) &= P(\{TTHT, TTHH\}) \\ &= P(\{TTHT\} \sqcup \{TTHH\}) \\ &= P(\{TTHT\}) \sqcup P(\{TTHH\}) \\ &= p(1-p)^3 + p^2(1-p)^2 \end{aligned}$$

11/10/14

HW 5 due Wed Nov 19
HW 6 due Wed Dec 3
Exam 3 Mon Dec 8

Last time we discussed Andrey Kolmogorov's (1933) axioms for probability.

Definition: A probability space is a pair (S, P) where S is a set called the "sample space" and P is a function

$$P : \wp(S) \longrightarrow \mathbb{R}$$

satisfying 3 axioms

1. For all $E \subseteq S$, $0 \leq P(E) \leq 1$.

2. $P(S) = 1$

3. For $E, F \subseteq S$ with $E \cap F = \emptyset$ we have

$$P(E \cup F) = P(E) + P(F).$$

Recall: Given sets A, B, C, we define the notation

$$\text{"}A = B \sqcup C\text{"} := \text{"}A = B \cup C \text{ and } \emptyset = B \cap C\text{"}.$$

The third axiom can be rewritten as

$$P(E \sqcup F) = P(E) + P(F)$$

If E and F are not disjoint we proved that the correct formula is:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

This is called the "Principle of Inclusion-Exclusion".

Very Special Case:

If the sample space S is finite, then there is an obvious choice for the probability function:

For each outcome $x \in S$ we define

$$P(\{x\}) := \frac{1}{\# S}$$

and we say that the outcomes are <u>equally likely</u>. In this case the probability of a general event $E \subseteq S$ is

$$P(E) = \frac{\# E}{\# S}.$$

Why? Suppose $E = \{x_1, x_2, \ldots, x_k\}$. Then we have a disjoint union

$$E = \{x_1\} \sqcup \{x_2\} \sqcup \cdots \sqcup \{x_k\}$$

and hence

$$P(E) = P(\{x_1\}) + P(\{x_2\}) + \cdots + P(\{x_k\})$$

$$= \frac{1}{\# S} + \frac{1}{\# S} + \cdots + \frac{1}{\# S}$$

$$= \frac{k}{\# S} = \frac{\# E}{\# S}. \quad ///$$

[Notation: I will write

$$P(x) := P(\{x\})$$

if you don't mind.]

Example: Roll 2 fair 6-sided dice.

Say one die is red and one is blue, so we can tell them apart. The sample space is

$$S = \{1,2,3,4,5,6\}^2$$

$$= \{(1,1), (1,2), \ldots, (1,6), \\ (2,1), (2,2), \ldots, (2,6), \\ \vdots \\ (6,1), (6,2), \ldots, (6,6)\}.$$

Note that $\# S = 36$. Since the dice are fair, each outcome is equally likely and we have

$$P(x) = \frac{1}{36} \quad \text{for all } x \in S.$$

What is the probability of "rolling a 6"?

We define

$E :=$ "rolling a 6"

$= \{(1,5), (2,4), (3,3), (4,2), (5,1)\}.$

So the probability is

$$P(E) = \frac{\#E}{\#S} = \frac{5}{36}.$$

On average we will "roll a 6" in 5 out of every 36 rolls of the dice.

What is the probability that the red die shows a ?

Let $F :=$ "red die shows 4"

$= \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}.$

Then $P(F) = \dfrac{\#F}{\#S} = \dfrac{6}{36} = \dfrac{1}{6}$

That makes sense.

What is the probability that "The sum of the dice is 6 OR the red die shows 4"?

This is the event $E \cup F$. We can compute $P(E \cup F)$ in two ways:

1. Count the outcomes.

$$E \cup F = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$$

$$\cup \{(4,1),\cancel{(4,2)},(4,3),(4,4),(4,5),(4,6)\}$$

$$= \{(1,5),(2,4),(3,3),(4,2),(5,1),(4,1),(4,3),(4,4),(4,5),(4,6)\}.$$

Hence

$$P(E \cup F) = \frac{\#(E \cup F)}{\#S} = \frac{10}{36}.$$

2. Use Inclusion-Exclusion

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

We already know $P(E) = \frac{5}{36}$ and $P(F) = \frac{6}{36}$.

To compute $P(E \cap F)$ note that

$E \cap F$ = "dice sum to 6 AND red die shows 4"
$= \{(4,2)\}$.

So $P(E \cap F) = \frac{\#(E \cap F)}{\# S} = \frac{1}{36}$.

Finally, we have

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

$$= \frac{5}{36} + \frac{6}{36} - \frac{1}{36}$$

$$= \frac{5+6-1}{36} = \frac{10}{36}.$$

Method 2 was easier because it involved less counting.

Example: Flip a biased coin n times
Assume that $P(H) = p$.

The sample space is

$$S = \{H, T\}^n, \text{ with } \#S = 2^n.$$

This time the events are <u>not</u> equally likely. If the sequence $x \in S$ contains k H's and $n-k$ T's, then

$$P(x) = p^k (1-p)^{n-k}.$$

This is not the same for every x.

What is the probability that we get at least 2 H's?

Let $E \subseteq S$ be the set of sequences with at least 2 H's.

$$P(E) = ?$$

Let's try a trick. We know that

$$P(E) = 1 - P(E^c)$$

where E^c is the set of sequences with 0 or 1 H's. These are fairly easy to describe:

$$E^c = \{ (T, T, T, \ldots, T), \\ (H, T, T, \ldots, T), \\ (T, H, T, \ldots, T), \\ \vdots \\ (T, T, \ldots, T, H) \}.$$

The probability is

$$P(E^c) = 1 \cdot (1-p)^n + n \cdot (p)^1 (1-p)^{n-1}$$

$$= (1-p)^{n-1} (1 + np).$$

Hence

$$P(E) = 1 - P(E^c)$$

$$= 1 - (1-p)^{n-1} (1 + np).$$

Remark: If the coin is fair ($p = \frac{1}{2}$) then the outcomes become equally likely.

$$\# S = \# \{H, T\}^n = 2^n$$

and for all $x \in S$ we have

$$P(x) = \frac{1}{2^n}.$$

What is the probability of getting exactly k heads?

Let E = sequences of length n with k H's and $n-k$ T's.

We know that $\# E = \binom{n}{k}$, and hence

$$P(E) = \frac{\# E}{\# S} = \frac{\binom{n}{k}}{2^n}$$

Does this agree with the general formula?

In general we have

$P(k\ H\text{'s in } n \text{ tosses})$

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

Put $p = \frac{1}{2}$ to get

$$\binom{n}{k} \left(\frac{1}{2}\right)^{k} \left(1 - \frac{1}{2}\right)^{n-k}$$

$$= \binom{n}{k} \left(\frac{1}{2}\right)^{\cancel{k}} \left(\frac{1}{2}\right)^{n-\cancel{k}}$$

$$= \binom{n}{k} \left(\frac{1}{2}\right)^{n}$$

$$= \frac{\binom{n}{k}}{2^n} .$$

Yes!

Example: Any experiment can be a "coin flip"

Roll 2 fair dice 24 times. What is the probability of getting "snake eyes" (i.e., (1,1)) at least once?

The probability of (1,1) in a single roll is 1/36. So the probability of getting (1,1) k times in 24 rolls is

$$\binom{24}{k}\left(\frac{1}{36}\right)^k\left(\frac{35}{36}\right)^{24-k}$$

(Each roll is a coin flip with $p = 1/36$.)
The probability of getting (1,1) at least once is

$$1 - P(\text{getting } (1,1) \text{ zero times})$$

$$= 1 - \binom{24}{0}\left(\frac{1}{36}\right)^0\left(\frac{35}{36}\right)^{24}$$

$$= 1 - \left(\frac{35}{36}\right)^{24} \approx 0.4914 = 49.14\%$$

Discussion:

This is one of the problems asked by the gambler de Méré to Pascal in 1640. He wanted to know which of the following is more likely:

1. At least one 1 in a single roll of 4 dice, or

2. At least one (1,1) in 24 rolls of a pair of dice.

He felt intuitively that 1. & 2. should be equally likely, but his gambling experience suggested that 1. happens more often.

Can you solve the problem?

11/12/14

HW 5 due Wed Nov 19
HW 6 due Wed Dec 3
Exam 3 Mon Dec 8

The Chevalier de Méré (1607–1684) was an amateur mathematician and "gentleman gambler". In 1640 he wanted to know which of the following is more likely:

1. Getting at least one 1 in a single roll of 4 fair dice.

2. Getting at least one (1,1) in 24 rolls of 2 fair dice.

He believed based on faulty mathematics that 1 & 2 are equally likely, but for some reason he won money on 1 and lost money on 2.

The (less ~~non~~ amateur) mathematicians Pascal and Fermat were able to explain why.

Game 1. Roll 4 fair dice once.

The sample space is

$$S = \{1,2,3,4,5,6\}^4.$$

All outcomes are equally likely:

$$P(x) = \frac{1}{\#S} = \frac{1}{6^4} \quad \forall\, x \in S.$$

Let E = "get at least one 1"
$= \{\ ?\ \}$
It's a big set.

Instead we'll consider the complement

E^c = "get NO 1's"
$= \{2,3,4,5,6\}^4.$

The probability of E^c is

$$P(E^c) = \frac{\#E^c}{\#S} = \frac{5^4}{6^4}.$$

Hence the probability of E is

$$P(E) = 1 - P(E^c)$$
$$= 1 - \frac{5^4}{6^4}$$
$$\approx 0.5177 = 51.8\%$$

Note that $P(E) > 50\%$ as observed experimentally be le Méré.

Game 2.

Roll 2 fair dice 24 times. We can think of each roll as a coin flip with

"heads" = "get (1,1)"
"tails" = "don't get (1,1)"

By separate analysis we compute

$$P(\text{heads}) = \frac{1}{36}, \quad P(\text{tails}) = \frac{35}{36}.$$

Since we roll the dice 24 times, the sample space is

$$S = \{H, T\}^{24}.$$

The outcomes are <u>not</u> equally likely, so we have to be careful. We are interested in the event

$$E = \text{"get at least one } H\text{"}$$
$$= \{\ ?\ \}$$

It's a big set.

Instead we'll consider the complement

$$E^c = \text{"get NO } H\text{'s"}$$
$$= \text{"get all } T\text{'s"}$$
$$= \{\underbrace{(T, T, T, T, \ldots, T)}_{24 \text{ times}}\}$$

Since $P(T) = \frac{35}{36}$ we have

$$P(E^c) = P((T, T, T, \ldots, T))$$
$$= P(T)^{24} = \left(\frac{35}{36}\right)^{24}.$$

Hence

$$P(E) = 1 - P(E^c)$$
$$= 1 - \left(\frac{35}{36}\right)^{24}$$
$$\approx 0.4914 = 49.14\%$$

Note that $P(E) < 50\%$, as observed by de Méré.

Now we return to the problem of "expected value".

Let S be a sample space with probability function

$$P : \wp(S) \longrightarrow \mathbb{R}$$

Q: In this case, what is the <u>expected</u> (or <u>average</u>) outcome of the experiment?

A: In general this question makes no sense!

Example: An urn contains 3 balls: one red, one blue, one green. You reach in and grab one at random. What is the expected outcome?

What should we say?

$$\frac{\text{red} + \text{blue} + \text{green}}{3} \; ?$$

This is nonsense!

We can only talk about expected value when the outcome is a number.

Example: Reach in and grab one ball. Record the number of red balls you get. Call it X.

$$P(X=1) = 1/3$$
$$P(X=0) = 2/3$$

The expected value of X is

$$E(X) = 0 \cdot P(X=0) + 1 \cdot P(X=1)$$

$$= 0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3} = \frac{1}{3}.$$

We expect to get $\frac{1}{3}$ red balls.

In general, given a probability space

$$P : \wp(S) \to \mathbb{R}$$

we define a random variable to be a function

$$X : S \to \mathbb{R}$$

that attaches a number $X(s) \in \mathbb{R}$ each outcome $s \in S$.

Example: Flip a biased coin 3 times.

$$S = \{H, T\}^3.$$

Let $X : S \to \mathbb{R}$ be the "number of heads", so

$$X((H,H,H)) = 3$$
$$X((H,H,T)) = 2$$
etc.

For each <u>number</u> k we define the event

$$E_k = \text{“} X = k \text{”}$$
$$= \{ s \in S : X(s) = k \}$$

In our example,

$$E_2 = \text{“} X = 2 \text{”}$$
$$= \{ (H,H,T), (H,T,H), (T,H,H) \}.$$

Finally we use Archimedes' "Law of the Lever" to define the <u>expected value</u> of the random variable X:

$$\boxed{E(X) := \sum_k k \cdot P(X=k)}$$

In our example (flip biased coin 3 times, with $X(s) = \#$ heads in s) we have

$$P(X=0) = (1-p)^3$$

$$P(X=1) = 3p(1-p)^2$$

$$P(X=2) = 3p^2(1-p)$$

$$P(X=3) = p^3.$$

So the expected number of heads is

$$E(X) = \sum_k k \cdot P(X=k)$$

$$= 0 \cdot (1-p)^3 + 1 \cdot 3p(1-p)^2 + 2 \cdot 3p^2(1-p) + 3 \cdot p^3$$

$$= 3p\left[(1-p)^2 + 2p(1-p) + p^2\right]$$

$$= 3p\left[1 - \cancel{2p} + \cancel{p^2} + \cancel{2p} - \cancel{2p^2} + \cancel{p^2}\right]$$

$= 3p[1] = 3p.$

Does that surprise you? It shouldn't.

Recall: Flip a biased coin n times and let X be the number of heads.

$$E(X) = \sum_k k \cdot P(X=k)$$

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

$$= np.$$

(We computed this previously.)

Another Example:

Suppose an urn contains R red and G green balls. You reach in and grab n balls. Let $X = \#$ red balls you get.

Compute $P(X=k)\ \forall k$.

The sample space is

$$S = \text{all subsets of } n \text{ balls}.$$

Note that $\#S = \binom{R+G}{n}$. Since all outcomes are equally likely we have

$$p(E) = \frac{\#E}{\#S} = \frac{\#E}{\binom{R+G}{n}}$$

for any event $E \subseteq S$. Now consider the event

$$E_k = \text{"} X = k \text{"}$$
$$= \{ \text{subsets of } n \text{ balls in which } k \text{ balls are red} \}.$$

We know from HW4.2 that

$$\#E_k = \binom{R}{k}\binom{G}{n-k}$$

Hence the probability of getting k red balls is

$$P(X=k) = P(E_k) = \frac{\# E_k}{\# S}$$

$$= \frac{\binom{R}{k}\binom{G}{n-k}}{\binom{R+G}{n}}.$$

There is a fancy name for this random variable. It is called the

"hypergeometric distribution"

Reality Check: Do we have

$$\sum_k P(X=k) = 1 \; ?$$

Let's see ...

$$\sum_k \frac{\binom{R}{k}\binom{G}{n-k}}{\binom{R+G}{n}} = 1 \; ?$$

$$\sum_k \binom{R}{k}\binom{G}{n-k} = \binom{R+G}{n} \; ?$$

Yes. This is a true identity called the "Vandermonde convolution". You proved it on HW 4, Problem 2.

Finally, let's consider the expected number of red balls.

On one hand we have

$$E(X) = \sum_k k \cdot P(X=k)$$

$$= \sum_k k \cdot \frac{\binom{R}{k}\binom{G}{n-k}}{\binom{R+G}{n}}$$

On the other hand, the ratio of red balls in the urn is

$$\frac{R}{R+G}.$$

So if we grab n we expect that

$n \cdot \frac{R}{R+G}$ of them will be red.

This leads us to guess that

$$\sum_k k \cdot \frac{\binom{R}{k}\binom{G}{n-k}}{\binom{R+G}{n}} = \frac{nR}{R+G}$$

$$\sum_k k \cdot \binom{R}{k}\binom{G}{n-k} = \frac{nR}{R+G}\binom{R+G}{n}$$

This is true!

But I don't feel like proving it.
I'm convinced by the inutive
argument.