期望值 I

随机变量的**期望**或**期望值**是一个单独数字，其告诉您很多关于变量的行为。大致上，期望是平均值， 这里的每个值根据其出现的概率来衡量出来。 正式地，在样本空间*S* *上的*定义的随机变量R 的期望值是：

$$\operatorname{Ex}(R) = \sum_{w \in S} R(w) \Pr(w)$$

为了赞赏它的重要性，假设*S*是在课堂上的学生，并且我们一致地随机选择一个学生。 令*R*是选择的学生的测验成绩。 然后Ex(R)正是课堂的平均值----在他们测试以后大家想要知道第一件事。同样，期望是你通常想要确定关于其中任一随机变量的第一件事。

我们来研究一个例子。 令 *R* 是有公平的、六个面的骰子的数字。 然后 *R* *的*期望值是：

$$\begin{aligned}
\operatorname{Ex}(R) &= \sum_{k=1}^{6} k\left(\frac{1}{6}\right) \\
&= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \\
&= \frac{7}{2}
\end{aligned}$$

这个演算说明，名字“期望值”是有一点引入歧途的; 任意变量也许从未实际上取那个值。 您不可能用一个普通的骰子滚出值3 • (1/2)的!

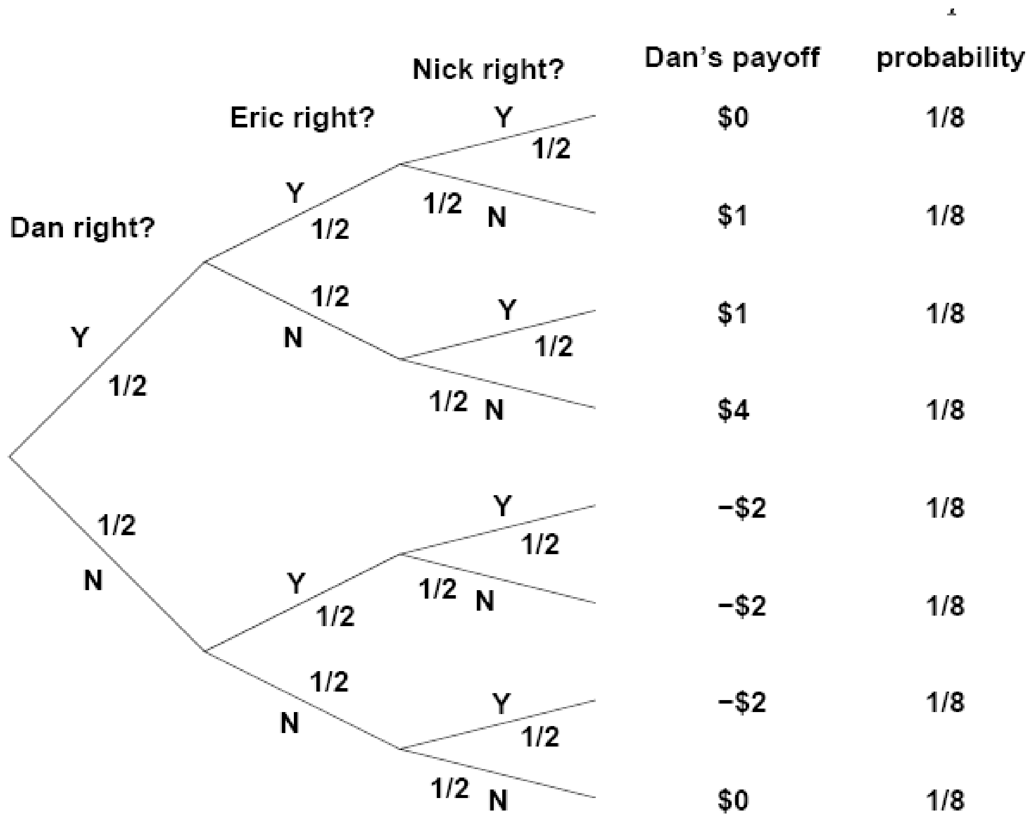
1 用硬币打赌

丹、埃里克和尼克决定玩一个有趣游戏。 每个游戏者在桌上放 $2，秘密地写下“正面”或“反面”。 然后他们中的一个抛一枚公平的硬币。 在桌上 $6 在正确地预测硬币抛的结果的游戏者之中均匀地分配。如果大家的猜测都不正确，

那么大家收回他们的钱。

在重复了多次游戏以后，丹已经失去了很多钱，比由坏运气给出的解释还要多。这是怎么回事？

为这个问题下面一个树图给出了解答，假定大家都有1/2 的概率猜对，且大家正确的猜测是独立的。



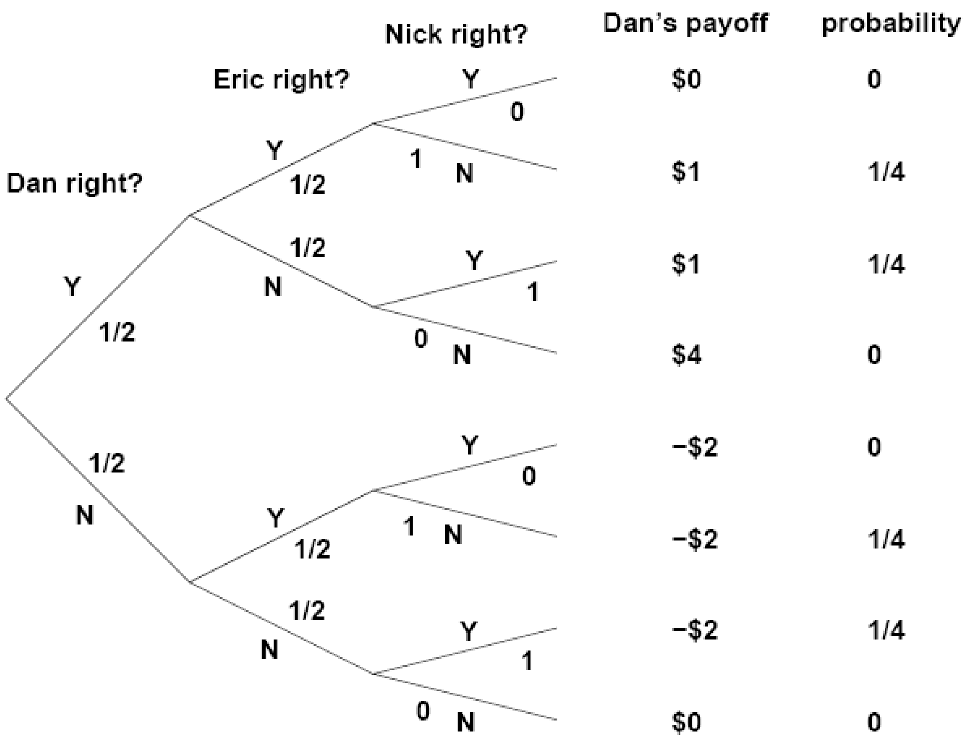
在 “payoff” 栏，我们解释了丹为了玩游戏，必须投入 $2 的事实。 例如，如此如果他正确地猜测并且埃里克和尼克错误，然后他拿走了全部在桌子上的$6 ，但他的净盈利仅是$4。 树形图上计算，丹的期望的支付是：

$$\begin{aligned}\mathrm{Ex}\,(\text{payoff}) &= 0\cdot\frac{1}{8}+1\cdot\frac{1}{8}+1\cdot\frac{1}{8}+4\cdot\frac{1}{8}+(-2)\cdot\frac{1}{8}+(-2)\cdot\frac{1}{8}+(-2)\cdot\frac{1}{8}+0\cdot\frac{1}{8}\\ &=0\end{aligned}$$

如此比赛完全是公平的! 随着时间的过去，他不应该赢钱和输钱。

诀窍是尼克和埃里克合作; 特别是，他们总做相反的猜测。 如此我们的假定大家是独立地正确的的是错误的; 实际上 事件尼克是正确的和埃里克是正确的互相排斥! 结果，丹从来就不能赢取在桌山的所有的钱。 当他猜测正确时，他必

须总和其他人分享他的赢的钱。 这降低他的整体期望，下面作为校正的树图说明：



从这个修改过的树图，我们可以得出丹的实际期望支付：

$$\begin{aligned}\mathrm{Ex}\,(\text{payoff}) &= 0\cdot 0+1\cdot\frac{1}{4}+1\cdot\frac{1}{4}+4\cdot 0+(-2)\cdot 0+(-2)\cdot\frac{1}{4}+(-2)\cdot\frac{1}{4}+0\cdot 0\\ &= -\frac{1}{2}\end{aligned}$$

因此平均每场比赛他失去半美元!

在许多打赌的游戏中有产生微妙欺诈相似的机会。例如，一个小组朋友曾经组织一个足球彩票，这里每个参加者会猜测相对传播的每周每场比赛的结果。这对您来说没有什么意义，但结果是每周大家在有效地在12或13 次抛硬币的的结果上打赌。正确地预言多数抛硬币的结果的会赢很多钱。 组织者认为，根据第一个树形图，从上到下发誓没有办法得到一个不公平的“边”。 但实际上参加者的数量是足够小的两个游戏者对立地打赌就可能获得实质的优势!

另一个例子涉及一位前MIT统计学教授， Herman Chernoff。
国家彩票是最坏的赌博的比赛，因为国家的支出，只是其收益的一小部分。 Chernoff 找到了获胜的方法! 这是一次典型的抽奖规则：

- 所有游戏者支付$1 参与游戏，从1到36选择4个数字。
- 国家从1 到 36，一致地随机得出4个数。
- 国家在正确地猜测的人民之中分出1/2 收集的钱，

并且使用另一半花费修理Big Dig(项目名称)。

这是很多象我们打赌的比赛，除了有更多的玩家和更多选择。 Chernoff 发现一小的数子集由大部分人口—许多人选择显然地认为同一个方式。 正如玩家球员合作输掉一样! 如果任何他们中的一个恰当地猜测，则他们也必须与许多其他参与者分享他们的所得。 通过一致地随机选择数字， Chernoff是不太可能得到这些偏爱的序列之一。 因此，如果他赢取了， 他可能得到所有的奖金! 通过分析实际状况抽奖数据，他确定这种方式他可能用一美元赢取平均7美分!

2 期望的等价定义

存在其他写作期望的定义的方式。 有时使用这些其他公式之一能使计算期望更加容易。 一个选择是一起编组取同一价值的随机变量的所有结果。

定理 1

$$\mathrm{Ex}(R) = \sum_{x \in \mathit{range}(R)} x \cdot \Pr(R = x)$$

证明： 我们将变换左边成右边。 令[R = x]是 R = x的事件。

$$\begin{aligned}
\mathrm{Ex}(R) &= \sum_{w \in S} R(w) \Pr(w) \\
&= \sum_{x \in \mathrm{range}(R)} \sum_{w \in [R=x]} R(w) \Pr(w) \\
&= \sum_{x \in \mathrm{range}(R)} \sum_{w \in [R=x]} x \Pr(w) \\
&= \sum_{x \in \mathrm{range}(R)} \left(x \cdot \sum_{w \in [R=x]} \Pr(w) \right) \\
&= \sum_{x \in \mathrm{range}(R)} x \cdot \Pr(R = x)
\end{aligned}$$

在第二条线，我们把总和分为二部分。 外面总和取所有的值随机变量x可能采取的值，并且内的总和在取那值的所有结果上取总和。 因此，我们仍然是正确地在样本空间的每个结果上求和一次。 在最后一行，我们使用事件[R = x]的概率的定义。

推论2 如果R是自然数值随机变量，那么：

$$\mathrm{Ex}(R) = \sum_{i=0}^{\infty} i \cdot \Pr(R = i)$$

有另一个写随机变量期望值的一种方式， N的期望值= {0,1,2，….
}
定理 3 如果R是一个自然数随机变量，那么：

$$\mathrm{Ex}(R) = \sum_{i=0}^{\infty} \Pr(R > i)$$

证明： 考虑总和：

$$\begin{array}{ccccccc}
\Pr(R=1) & + & \Pr(R=2) & + & \Pr(R=3) & + & \cdots \\
 & + & \Pr(R=2) & + & \Pr(R=3) & + & \cdots \\
 & & & + & \Pr(R=3) & + & \cdots \\
 & & & & & + & \cdots
\end{array}$$

列的总和到1·Pr (R= 1)， 2·Pr (R= 2)， 3·Pr (R= 3)，等等。 因此，整体总和是相等的对：

$$\sum_{i=0}^{\infty} i \cdot \Pr(R = i) = \mathrm{Ex}(R)$$

这里，我们使用推论2。另一方面，对Pr (R>0)， Pr (R>1)， Pr (R>2)等等。 因此，整体总和也是也等于：

$$\sum_{i=0}^{\infty} \Pr(R > i)$$

这两个表达式的整体总和的一定是相等的，证明了定理。

2.1对失败的平均时间

让我们看看这些供选择的期望值的定义之一是特别有用的问题。 如果它没有崩溃，计算机程序在每个小时后以概率p崩溃。 什么是期望时间，直到程序崩溃？

如果让我们是几小时的数量直到崩溃，然后对我们的问题的答复E(R)。 这是自然数值的变量，因此我们可能使用公式：

$$\operatorname{Ex}(R) = \sum_{i=0}^{\infty} \Pr(R > i)$$

我们有R > i，只有当系统在i机会崩快依然是稳定的，这发生的概率是 $(1-p)^i$。把这个带入上面给出的公式：

$$\begin{aligned}\operatorname{Ex}(R) &= \sum_{i=0}^{\infty} (1-p)^i \\ &= \frac{1}{1-(1-p)} \\ &= \frac{1}{p}\end{aligned}$$

在第二行的闭形来自无限几何级数和的闭形，这里连续项的比是1-p。

因此，例如，如果有1%机会节目在每个小时的结尾的时候崩快，那么直到节目崩溃时候的期望时间，是1/0.01 =100小时。这里的通用院子是容易记忆的：如果系统在每步概率为p失败，那么到第一个失败的期望的步的数字是1/p。

2.2 生女孩

一对夫妇真正地想要有女婴。 有50%机会他们有个女孩，并且他们的孩子性别是相互独立的。 如果夫妇坚持直到他们得到女孩的时候才要孩子，则他们应

该首先等多少个男婴？

这是一个真正地是早先问题的变种。 问题， “多少个小时后，程序崩溃？” 数学上是相同的问题， “夫妇必须有多少个孩子后，他们才能得到女孩？”在这种情况下，崩溃对应于有女孩，因此我们应该设置p = 1/2。 由在先的分析，夫妇应该期望有的女婴1/p =2孩子以后，他们应该期望有一个男孩。

3 线性期望

期望值服从一个美妙的规则称为线性期望。 这说总和的期望期望的总和。

定理 4(期望的线性性) 对每个随机变量R_1和R_2 ：

$$\mathrm{Ex}(R_1 + R_2) = \mathrm{Ex}(R_1) + \mathrm{Ex}(R_2)$$

证明：令S是样本空间。

$$\begin{aligned}
\mathrm{Ex}(R_1 + R_2) &= \sum_{w \in S} (R_1(w) + R_2(w)) \cdot \Pr(w) \\
&= \sum_{w \in S} R_1(w) \cdot \Pr(w) + \sum_{w \in S} R_2(w) \cdot \Pr(w) \\
&= \mathrm{Ex}(R_1) + \mathrm{Ex}(R_2)
\end{aligned}$$

期望线性推断由归纳可得一般化到任意有限的随机变量的集合：

推论5 对任何随机变量R_1， R_2，…， R_k，

$$\mathrm{Ex}(R_1 + R_2 + \cdots + R_k) = \mathrm{Ex}(R_1) + \mathrm{Ex}(R_2) + \cdots + \mathrm{Ex}(R_k)$$

期望线性性是很美妙的原因在于，不同于许多其他概率规则，随机变量没有要求是独立的。 这现在大概听起来“yeah，管它的呢”术语。 但是，当您使用期望线性给出分析，某人将错误几乎不变地说， “不，您错了。 有各种各样复杂的您忽略的依赖.”，但是那是基本的期望线性魔术： 您能忽略这样依赖!

3.1二个骰子的期望值

两个公平的骰子的期望值是什么？令随机变量R_1是在第一个骰子的数字，并且令R_2是在第二骰子的数字。 在这些笔记的开始，我们说明，一个骰子的期望值$3\frac{1}{2}$ 。使用期望的线性，我们可以找到总和的期望值：

$$
\begin{aligned}
\operatorname{Ex}(R_1 + R_2) &= \operatorname{Ex}(R_1) + \operatorname{Ex}(R_2) \\
&= 3\frac{1}{2} + 3\frac{1}{2} \\
&= 7
\end{aligned}
$$

注意我们没有必须假设，二个骰子是独立的。 二个骰子的期望的总和是7，即使他们一起被胶合! (这是，在莫名其妙条件下地胶合
不改变各自的骰子不公平的权重。)

为一棵树提供期望的总和是7是难的; 有36个种条件。 并且，如果我们没有假设骰子是独立的，工作就是恶梦!

3.2 帽子检测问题

有一个n人检查他们的帽子的晚餐会。 在晚餐期间，帽子被混合，因此每个人之后接受一个任意帽子。 特别是，每个人得到他的自己的帽子的概率是1/n.。他们自己得到帽子人的期望的数量是什么？

没有线性期望，这是回答的一个非常困难的问题。 我们也许尝试以下。 令随机变量R是得到他们自己的帽子人的数量。 我们想要计算E(R)。 由期望的定义得，我们有：

$$
\operatorname{Ex}(R) = \sum_{k=0}^{\infty} k \cdot \Pr(R = k)
$$

现在我们麻烦了，因为求值Pr (R = k)是混乱得和我们然后需要把这混乱带入总和。 此外，有所有希望，我们会需要固定帽子的每排列的概率。 例如，我们也许假设，帽子的所有组合是等概率得。

让我们尝试使用期望得线性的尝试。 前面，令随机变量R是得到他们自己的帽子人的数量。 技巧是表达R作为指示变量的一个总和。 特别是，令 R_i是显示第i个人得到他自己的帽子的事件。 即R_i =1是他得到他自己的帽子的事件和R_i =0是他得到错误帽子的事件。 得到他们自己的帽子人的数量是这些指示器的总和：

$$
R = R_1 + R_2 + \cdots + R_n
$$

这些显示变量不是相互独立的。 例如，如果n −1人全部得到他们自己的帽子，那么最后人肯定接受他自己的帽子。 但是，因为我们计划使用期望线性，我们不必考虑独立性!

让我们取上面等式的两边的期望值并且应用期望的线性：

$$\begin{aligned}\mathrm{Ex}(R) &= \mathrm{Ex}(R_1 + R_2 + \cdots + R_n) \\ &= \mathrm{Ex}(R_1) + \mathrm{Ex}(R_2) + \cdots + \mathrm{Ex}(R_n)\end{aligned}$$

所有剩下的就是计算显示变量R_i的期望值。我们将用基本的事实，其在它子集一边是值得记忆的：

事实 1 指示随机变量的期望值是指示器是1的概率。以符号：

$$\mathrm{Ex}(I) = \Pr(I = 1)$$

证明：

$$\begin{aligned}\mathrm{Ex}(I) &= 1 \cdot \Pr(I = 1) + 0 \cdot \Pr(I = 0) \\ &= \Pr(I = 1)\end{aligned}$$

因此，我们现在需要只计算 Pr (R_i =1)，是第 i 个人得到他自己的帽子的概率。因为每个人是可能得到一个像别人的帽子的，这正是 1/n。把所有的放到一起，我们有：

$$\begin{aligned}\mathrm{Ex}(R) &= \mathrm{Ex}(R_1) + \mathrm{Ex}(R_2) + \cdots + \mathrm{Ex}(R_n) \\ &= \Pr(R_1 = 1) + \Pr(R_2 = 1) + \cdots + \Pr(R_n = 1) \\ &= n \cdot \frac{1}{n} = 1.\end{aligned}$$

因此我们应该盼望 1 个人平均得到他自己的帽子。

注意到，我们没有假设所有排列是等概率的或者甚至所有的排列是可能的。我们只需要知道每个人用概率 1/n.接受了他的自己的帽子。 这让我们的解很一般化，正如下一个例子说说明的。

3.3 中国开胃菜问题

有 n 人在中国餐馆的一张圆桌坐着。 在桌子上，有大懒惰苏珊安排的 n 不同的开胃菜。 每个人开始用力嚼直接地在他或她前面开胃菜。 然后某人转动懒惰苏珊，以便大家面对一道任意开胃菜。最终获得最初食用开胃菜的期望的什么？

这是检测帽子的一个特殊情况。用在帽子位置替换开胃菜。 在帽子检测问题种，我们假设每个人用概率 1/n 接受了他的自己的帽子。 在那之外，我们没有做关于帽子怎样可能被排列的假定。 这个问题是一个特殊情况，因为我们知道开胃菜周期地被转移相对他们最初的位置。 这意味着或者大家让他们原始的开胃菜回来，或者没人做。 但是我们的早先仍然分析成立： 得到他们自己的开胃菜回来的人数的期望的是 1。