Math

一、概率

主要是要深入理解先验和后验，理清概率的来源，了解测度和概型，常用的有贝叶斯概率和全概率公式。

1.最大的麦穗

答：1/e = 36.7% ， 前面的36.7%的麦穗放过，只记录一个最大值。接下来的麦穗只要大于前面的最大值，就拾起。可参考https://www.zhihu.com/question/66465943

2.抛硬币连续k次正面向上

答：对于独立重复实验，不断实验，直到某一个事件连续发生k次。设该事件在每次实验中发生的概率都是p，设实验次数的期望是E(N_k)。那么从状态转移的思想考虑，可以得到：

$$E(N_k|N_{k-1}) = N_{k-1} + p \times 1 + (1-p) \times [1 + E(N_k)]$$

即在前面已经连续发生k-1次的情况下，有p的概率在这一次还是发生，则实验终结，而有1-p的概率不发生该事件，此时的状态回到最开始没有实验的状态，把它们的次数乘以概率加起来作为期望次数，再加上前面的连续k-1次发生该事件所用的实验次数，就得到了等式的左边。

接下来使用全概率公式，两边取期望，可以得到：

$$E(N_k) = E[E(N_k|N_{k-1})] = E(N_{k-1}) + p \times 1 + (1-p) \times [1 + E(N_k)]$$

从而有：

$$E(N_k) = \frac{1}{p} + \frac{E(N_{k-1})}{p}$$

迭代下去，可以解得：

$$E(N_k) = \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \ldots + \frac{1}{p^k}$$

不妨设这里求抛硬币连续4次正面向上，那么p=1/2，k=4，从而可得期望次数是30次。

可参考https://zhuanlan.zhihu.com/p/68358814。

3.圆上任取三个点组成锐角三角形

答：概率为1/4，可参考https://zhuanlan.zhihu.com/p/69530841。

4.生男生女比例

一个国家的人们都喜欢生男孩，如果生的是女孩，则会继续生直到生出男孩位置，如果生的是男孩，则停止。问这个国家的男女比例。

答：生男孩的平均期望是E(boy)=1，因为会一直生到男孩为止。或者构造方程 E(boy)=1/2 * 1 + 1/2 * E(boy)，同样有E(boy) = 1。现在考虑E(girl)。 考虑第一次生育，分为两种情况：如果是男孩，则男孩个数为1，女孩个数为0; 如果是女孩，则男孩个数为0，女孩个数为1。注意对于第二种情况，它的状态并未发生改变，即和还没有开始生育时没差别，该家庭生育女孩的期望依然是E(girl)。从而有E(girl)=1/2 * 0 + 1/2 * (E(girl) + 1), 即E(girl)=1。 因为任何家庭的地位都是等价的，且男女期望都是1，故比例是1:1。

5.抛硬币谁先抛到正面谁就赢

甲乙抛硬币比赛，甲先抛。谁先抛到正面谁赢。甲赢的概率是多少。 答：考虑第一轮，甲乙抛的结果有四种：正正，正反，反正，反反。甲赢的概率是2/4=1/2。如果是反正，比赛结束，比如是反反，比赛继续，由于事件独立，此时状态回到最开始未比赛的状态。

所以设甲赢的概率是X，可以得到X=1/2 + 1/4*X，故X=2/3，即甲赢的概率是2/3。此外，这种题还可以用等比数列求，然后取极限。

6.木棒截成三截组成三角形

答：设第一段截 x，第二段截 y，第三段 1 - x - y。首先有 0 < x < 1，0 < y < 1，0 < 1 - x - y < 1。画图可知，(x, y) 必须在单位正方形的左下角的半个直角三角形里，面积为 1 / 2。

其次，要组成三角形，要求两边之和大于第三边：x + y > 1 - x - y；x + 1 - x - y > y；y + 1 - x - y > x。从而0 < x < 1/2，0 < y < 1/2，1/2 < x + y < 1。画图可知，此时 (x, y) 必须在边长为 1/2 的三角形的右上角的半个直角三角形里，面积为 1/8。于是最终概率为 (1/8) / (1/2) = 1/4。

7.三角形上三只蚂蚁不相撞的概率

三角形的三个端点上有三只蚂蚁，蚂蚁可以绕任意边走，蚂蚁不相撞的概率是多少？

答：乍一看好像和三角形三边长度有关系，事实上是无关的。

首先，每个蚂蚁在方向的选择上有且只有 2 种可能，共有 3 只蚂蚁，所以共有 2 的 3 次方种可能，而不相撞有有 2 种可能，即全为顺时针方向或全为逆时针方向。故不相撞概率 = 不相撞 / 全部 = 2/8 = 1/4。

8.扑克牌里的大小王

一副扑克牌54张，现分成3等份每份18张，问大小王出现在同一份中的概率是多少？ 答：54张牌分成3等份，共有M=C(54,18) * C(36,18) * C(18,18)种分法。

其中大小王在同一份的分法有N= C(3,1) * C(52,16) * C(36,18) * C(18,18)种。因此所求概率为P=N/M=17/53。

9.宝剑升级

有一把宝剑，每使用一个宝石，有50%的概率会成功让宝剑升一级，50%的概率会失败。如果宝剑的级数大于等于5的话，那么失败会使得宝剑降1级。如果宝剑的级数小于5的话，失败没有效果。问题是：期望用多少个宝石可以让一把1级的宝剑升到9级？

答：定义Xi表示i级升级到i+1级所需的宝石个数，则EXi表示i级升级所需的期望宝石数。设总宝石数为X，则EX=EX1+EX2+...+EX8。显然EX1=EX2=EX3=EX4=2 EX5=1/2 * 1 + 1/2 * (1 + EX4 + EX5),得到EX5=4； EX6=1/2 * 1 + 1/2 * (1 + EX5 + EX6),得到EX6=6； EX7=1/2 * 1 + 1/2 * (1 + EX6 + EX7),得到EX7=8； EX8=1/2 * 1 + 1/2 * (1 + EX7 + EX8),得到EX8=10； 从而求和得到EX=36。

10.三门问题

可参考https://blog.csdn.net/qq_27782503/article/details/95952107与https://www.zhihu.com/topic/20046051/top-answers，答案是要换。

11.随机数生成

已知一随机发生器，产生 0 的概率是 p，产生 1 的概率是 1-p，现在要你构造一个发生器，使得它产生 0 和 1 的概率均为 1/2。

答：连续产生两个数，其组合以及概率如下：00 : p^2; 01 : p(1-p); 10 : (1-p) p; 11 : (1-p)^2。可以发现 01 和 10 组合的概率是相等的，只需要将其分别映射到 0 和 1 即可。即每次随机产生两个数，如果组合为00或11则丢弃，若为 01 则映射到 1，若为 10 则映射到 0，这样一来产生 0 和 1 的概率均为 1/2。

另一道相似的题目：已知一随机发生器，产生的数字的分布不清楚，现在要你构造一个发生器，使得它产生 0 和 1 的概率均为 1/2。

答：使用该随机发生器产生随机数 a，b，有以下 3 种情况：a < b；a == b；a > b。其中情况1和3是对称的，发生的概率相等，只需要将这两种情况分别映射到 0 和 1 即可，而遇到 a == b 时忽略。

12. 使用rand7()

已知有个 rand7() 的函数，随机返回1到7这七个自然数，怎样利用这个 rand7() 构造 rand10()，随机 1 ~ 10。

答：产生随机数的主要原则是**每个数出现的概率是相等的**，如果可以得到一组等概率出现的数字，那么就可以从中找到映射为 1 到 10 的方法。

rand7() 返回 1 ~ 7 的自然数，构造新的函数 (rand7() - 1) * 7 + rand7()，这个函数会随机产生 1 到 49 的自然数。

原因是 1 到 49 中的每个数只有唯一的第一个 rand7() 的值和第二个 rand7() 的值才能表示，于是它们出现的概率相等。但是这里的数字太多，可以丢弃 41 到 49 的数字，把 1 到 40 的数字分成 10 组，每组映射成 1 到 10 中的一个，于是可以得到随机的结果。

具体方法是，利用 (rand7() - 1) * 7 + rand7() 产生随机数 x，如果大于 40 则继续随机直到小于等于 40 为止，如果小于等于 40，则产生的随机数为 (x - 1) / 4 + 1。

补充：一个基本的想法是这样的，假设要用RandA()构造RandB()，如果A>B，非常容易解决，即大于B的数抛弃，继续生成下一个在1到B之间的数，包括B本身，很容易证明其均匀性。

如果A<B，则构造RandA2 = (RandA - 1) * A + RandA，表示生成从1到A2随机数的函数。如果A2仍然小于B，则继续构造RandA3 = (RandA2 - 1) * A2 + RandA2，知道构造出RandAn，而An大于B，这样的话，就回到刚才直接删减元素的方法即可。这种方法能保证其概率均匀性。

此外，扩展的问题有：利用随机生成a到b范围的数的随机函数 randAB来生成randCD，一样也是线性构造，不过复杂点。

13.涂球的期望时间

一个木桶里面有M个白球，每分钟从桶中随机取出一个球涂成红色（无论白或红都涂红）再放回，问将桶中球全部涂红的期望时间是多少

参考：设桶中有i个红球后再把全部球涂红的期望时间是T[i]，此时取出一个球。如果是红色的，概率是i/M，涂红后继续放回，此时的状态与开始未取球时相同，故期望时间仍为T[i]。如果是白色的，概率是1- (i/M)，涂红后放回，此时的期望时间是T[i+1]。从而可以得到：T[i] = i/M * (1+T[i]) + (1- (i/M)) * (1+T[i+1])。

从而有T[i] = T[i+1] + (M/(M-i))。又显然T[M]=0，所以T[0] = M/M + M/(M-1) + ... + M/1。

14.选球游戏

有50个红色弹球和50个蓝色弹球，分配在两个罐子里，随机选出一个罐子然后从里面随机选出一个弹球，怎么给出红色弹球最大的选中机会? 在你的计划里，得到红球的几率是多少？

参考：先考虑特殊情况，一个罐子全空，那么概率是P=1/2 * 0 +1/2 * 1/2 = 1/4。

设左边罐是R1+B1，大于0，右边罐子是R2+B2，也大于0。

那么拿到红球的概率是：P= 1/2 * (R1/(R1+B1)) + 1/2 * (R2/(R2+B2))。

将1/2提出来，考虑R1/(R1+B1) 和 R2/(R2+B2) 这两项，要么两项都是1/2，要么一项大于1/2一项小于1/2。

前者，最后的答案是1/2*(1/2 + 1/2) = 1/2。

后者，大于1/2的最大可能是1，小于1/2的最大可能是49/99。而这两者可以同时实现，即一个罐里面只放1个红球，另一个里面放49个红球和50个蓝球。概率为(1+49/99)/2 = 0.7475。

15.抛骰子期望

抛一个六面的色子，连续抛直到抛到6为止，问抛的次数的期望是多少？

参考：设次数为X，第一次就抛到6的概率是1/6。抛到其他的是5/6，需要重新抛，由于事件独立，所以状态回到刚开始未抛的时候，从而有：E(X) = 1/6 + 5/6 * (1+E(X))，从而E(X) = 6。

16.圆内随机取一点

在半径为1的圆中随机选取一点。

参考：在[0, 2pi)间随机选取一个角度，然后在[0, R]中选取一个长度，即可以确定一个点。**选长度的时候，选取的概率要和离圆心的距离成正比**。可以参考这个链接。

其他的有：

1. 如何在一个平行四边形内随机选一点?

   两条边AB和AC上各选一点E和F，然后构造平行四边形AEFG，G即为随机选取的在平行四边形之内的点

2. 如何在一个等腰三角形内随机选取一点?

   先在平行四边形内选取一点，如果超过对角线，返回与对角线对称点即可。

3. 确定角度后，如何按长度概率返回点?

   确定角度后，可以看做在一个很窄的三角形内选取一点，这时两条边平行，选取点位置就是两次选取[0, 1]内随机值之和，超过1再折回即可。可以这么想，将一个平行四边形ABCD的AB和AC边(AB=AC)捏在一起，就成了一个长为2*|AB|的直条。

   Python代码如下：

```python
from random import *
import math

def randPointInCircle():
    theta = random() * 2 * math.pi
    l = random() + random()
    l = l if l <= 1 else 2 - l
    return (l * math.cos(theta), l * math.sin(theta))
```

17.一个圆上任意取两点，求弦长的期望

参考：任取一点A，连接圆心O，然后再任取一点B，设角AOB的角度为theta，那么在某一点出现的概率是d(theta)/(2*pi)，可以认为弦长是不变的。用余弦定理，可以得到对应的弦长是sqrt(2-2 * cos(theta))。从而可以从0到2 * pi积分，即为其期望，为4/pi。

18.圆内任意两点间距离的期望

从一个圆内任取两点，求着两个点距离的期望值，其中圆内点的选取服从均匀分布。

参考：https://www.zhihu.com/question/279327736，两个点用极坐标表示，用余弦定理算距离，然后积分。

19.等车时间期望

1.一个圆上有一个站点P，有人过来等车，车的速度是2分钟每圈，人过来后等车的时间的期望。

2.如果车有两种速度，每经过站点P，有40%的概率切换到另一个速度，那么人过来等车的时间的期望又是多少。

3.有两班公交车，发车间隔时间分别为5分钟和10分钟，等到任意一班公交车均上车，求在车站等车时间的期望。

参考：**1.**车在圆上符合均匀分布，考虑均匀分布U(a,b)的期望是E= (a+b)/2，方差为D=(b-a)^2/12。所以很容易得到车最开始与人的距离的期望是圆周的一半，故等车时间的期望是2*1/2 = 1 分钟。

2.因为两种速度没有其他先验知识，可以认为最开始的速度为v1或者为v2的概率是一样的，即皆为1/2。其速度切换后概率也是1/2 : 1/2。即两种速度的概率对人来说都是1/2。

从另一个角度来看，v1和v2的地位是等价的，信息也是等价的，故p(v1)=p(v2)=1/2。

设等车的时间是X，车最开始离人的距离是Y，从而有E(X) = 1/2 * (E(Y) / v1) + 1/2 * (E(Y) / v2)。

又Y在0到2 * PI * R上均匀分布，从而E(Y) = PI * R。 从而E(X) = PI * R * (v1 + v2) / (2 * v1 * v2)。

3.设概率分布函数F(x)表示在x分钟内等到公交车的概率，那么:

$$F(x) = 1 - (1 - \frac{x}{5})(1 - \frac{x}{10}),\ x \in [0, 5]$$

对其求导得到概率密度函数f(x):

$$f(x) = \frac{3}{10} - \frac{x}{25}$$

则可以得到等待时间x的期望是：

$$E(x) = \int_0^5 xf(x)dx = \int_0^5 \frac{3}{10}x - \frac{x^2}{25}dx = \frac{25}{12}$$

注：这种题目也可以用几何概型算，设X坐标表示A车到达时间，Y坐标表示B车到达时间，画出积分区域，可以得到平均的等待时间是：

$$\iint_{0 \leqslant x \leqslant 5, 0 \leqslant y \leqslant 5} \frac{min(x, y))}{5 \times 10} dxdy = \frac{25}{12}$$

但是要注意的是，这道题**发车的间隔是固定的**，每5分钟发一辆车，在https://www.zhihu.com/questi on/36348471中，提出的问题是间隔为0到15分钟，而不是15分钟，即**发车间隔是随机**的。在下面的解 答中，正确的答案是3.40625分钟，而不是5.635分钟。具体的解释可以参考回答1和回答2。此外也可 以参考这篇博客。

二、进制和信息论

三、排列组合

1. x+y+z+m=10，x,y,z,m均为正整数, 取值组合的种数

可参考：https://blog.csdn.net/gmh77/article/details/77200658与https://zhuanlan.zhihu.com/p/33841038

答：这种排列组合问题有隔板法、递推法、绑定法等。这里使用隔板法，为C(n-1,m-1)。

10个球分成4堆，每一堆都至少有一个，所以只有中间的9个间隔插入3个隔板。从而C(9,3)=84。

2.将10个小球分到三个桶中，第三个桶里面的球的数量为偶数的概率

这三个桶依次排开，桶里面的球的数量可以为0到10，0也视为偶数，球没有区别。

答：由于桶可以为空的，不能直接用组合里面的插板法。要使每个桶里面非空，先往每个桶里面放一个球，那么题目就转化为用13个球放到3个依次排开的桶中，每个桶里面至少有一个球，第三个桶里面的球的数量为奇数的概率。此时可以用插板法，即13个球构造了12个空，从里面选2个空插入插板，从而分成三堆，情况数有

$$C_{12}^{2}=\frac{12!}{10!2!}=11*6=66$$

由于每个桶里面至少有一个球，那么第三个桶里面可能存在的球数为1到11。那么现在需要求出第三个桶里面的球数为1、3、5、7、9、11的情况一共有多少种。考虑第三个桶里面有x个球，那么前面有13-x个球放在两个桶中，继续用插板法处理即可，从而可以得到满足条件的情况数有：

$$C_{11}^{1}+C_{9}^{1}+C_{7}^{1}+C_{5}^{1}+C_{3}^{1}+C_{1}^{1}=11+9+7+5+3+1=36$$

所以所求概率为P= 36/66 = 6/11。

四、其他

1.米和沙

A容器中有4L沙子，B容器中有4L米，假设米和沙子密度一样。有一个C容器是300ml, 第一步，用C从A中舀300ml到B中，混合均匀，第二步，用C从B冲舀300ml到A中**混合均匀**。再重复第一步和第二步，问这四步之后，A中的米和B中的沙子谁多？

答：两步作为一轮，那么不管重复多少轮，A中损失的沙子的体积设为x，都被B得到了，那么B由于量不变，必然要失去x体积的米 。比较严谨的证明如下：A中原来有4a，B中原来有4b，一轮后，由于体积维持在4，所以可以设A中有xa和(4-x)b，B中有ya和(4-y)b，那么由于物质守恒，xa+ya=4a，所以x+y=4，所以4-x=y，即A中的米和B中的沙一样多。

待补充

1. https://zhuanlan.zhihu.com/p/42473598
2. https://github.com/arkingc/note/blob/master/interview/%E6%95%B0%E5%AD%A6%E6%99%BA%E5%8A%9B%E9%A2%98.md