HW2 yp2201 @ nyu.edu

1. Given $\tilde{a}$ a geometric random variable with parameter α,
geometric distribution is: $P_{\tilde{a}}(a) = (1-\alpha)^{a-1}\alpha^1$

if we assume $a_1 = \tilde{a}$ equals a for $a = 1,2,3,\dots$,
and $a_2 = \tilde{a}$ equals a for $a > 5$,

the probability should be $P_{\tilde{a}}(a_1) / P_{\tilde{a}}(a_2) = \dfrac{P_{\tilde{a}}(a_1 \cap a_2)}{P_{\tilde{a}}(a_2)}$

if) $a = 1,2,3,4,5$ then $P_{\tilde{a}}(a_1 \cap a_2) = 0$,
Since there will be no intersection of value a.

if) $a > 5$, then $P_{\tilde{a}}(a_1 \cap a_2) = P_{\tilde{a}}(a_1)$
Since every a_1 will be an intersection of a_2

$P_{\tilde{a}}(a_1) = (1-\alpha)^{(a_1-1)} \cdot \alpha^1$.

$$P_{\tilde{a}}(a_2) = P_{\tilde{a}}(a>5) = 1 - \left(P_{\tilde{a}}(1) + P_{\tilde{a}}(2) + P_{\tilde{a}}(3) + P_{\tilde{a}}(4) + P_{\tilde{a}}(5)\right)$$
$$= 1 - (\alpha + (1-\alpha)\cdot\alpha + (1-\alpha)^2\alpha + (1-\alpha)^3\alpha + (1-\alpha)^4\alpha)$$
$$= (1 - \alpha - (1-\alpha)\alpha - (1-\alpha)^2\alpha - (1-\alpha)^3\alpha - (1-\alpha)^4\alpha)$$
$$= ((1-\alpha)(1 - \alpha - (1-\alpha)\alpha - (1-\alpha)^2\alpha - (1-\alpha)^3\alpha)$$
$$= (1-\alpha)^2(1 - \alpha - (1-\alpha)\alpha - (1-\alpha)^2\alpha)$$
$$= (1-\alpha)^3(1 - \alpha - (1-\alpha)\alpha)$$
$$= (1-\alpha)^4(1-\alpha)$$
$$= (1-\alpha)^5$$

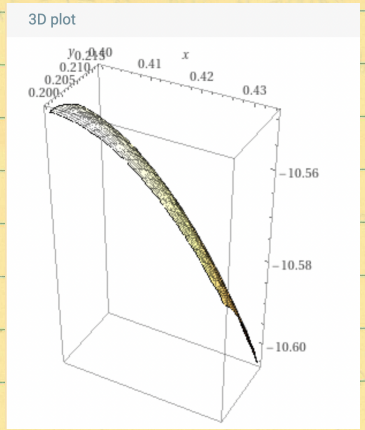
$\therefore$ if $a \le 5$, $P_{\tilde{a}}(a_1) / P_{\tilde{a}}(a_2) = \dfrac{P_{\tilde{a}}(a_1 \cap a_2)}{P_{\tilde{a}}(a_2)} = 0$

if $a > 5$, $P_{\tilde{a}}(a_1) / P_{\tilde{a}}(a_2) = \dfrac{P_{\tilde{a}}(a_1 \cap a_2)}{P_{\tilde{a}}(a_2)} = \dfrac{P_{\tilde{a}}(a_1)}{P_{\tilde{a}}(a_2)} = \dfrac{(1-\alpha)^{(a_1-1)}\cdot\alpha^1}{(1-\alpha)^5} = \boxed{(1-\alpha)^{(a_1-6)}\cdot\alpha^1}$

※ 2. (a) $P(\theta, a) = \frac{10!}{4!4!2!}\theta^4 a^2 (1-\theta-a)^4$

$$\log P(\theta, a) = \log\frac{10!}{4!4!2!} + 4\log(\theta) + 2\log(a) + 4\log(1-\theta-a)$$

<Plot>



(b) By taking the derivative of both θ and a, we can calculate θ, a.

1) $\frac{d}{d\theta}\log P(\theta, a) = 0 + \frac{4}{\theta} + 0 + \frac{4\times(-1)}{(1-\theta-a)} = 0$

$\Rightarrow \frac{4}{\theta} = \frac{4}{(1-\theta-a)}$, $\theta = (1-\theta-a)$ $\therefore \theta = \frac{1-a}{2}$

2) $\frac{d}{da}\log P(\theta, a) = 0 + 0 + \frac{2}{a} + \frac{4(-1)}{(1-\theta-a)} = 0$, $2a = 1-\theta-a$ $\therefore \theta = 1-3a$

Since $\theta = \frac{1-a}{2} = 1-3a$, $1-a = 2-6a$, $5a = 1$, $\boxed{a = \frac{1}{5}}$, $\boxed{\theta = \frac{1-\frac{1}{5}}{2} = \frac{2}{5}}$

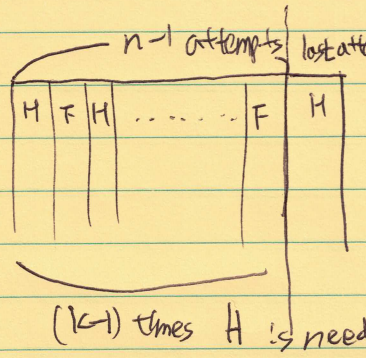
(c) $P(\text{Garry wins}) = \frac{4}{10} = \frac{2}{5}$

$P(\text{Anish wins}) = \frac{2}{10} = \frac{1}{5}$

$P(\text{draw}) = \frac{4}{10} = \frac{2}{5}$.

from 10 chess games.
This result is same as (b), which makes this nonparametric model same as parametric model

3. If a player needs total n attempts and hits a certain number k times, last attempt should be hitting a certain number.



(H : Hits a certain number)
F : fails to hit a certain number)

So, for $(n-1)$ attempts, we have $(k-1)$ Hits and $(n-1)-(k-1)$ none hits in binominal distribution.

Assuming the probability of success in each attempt as θ, and number of required attempts k, pmf should be as below

pmf) $\underline{{}_{n-1}C_{k-1} \cdot (\theta)^{k-1} \cdot (1-\theta)^{(n-1)-(k-1)}} \cdot \underbrace{\theta}_{\uparrow}$

$p(\text{last attempt} = \theta)$

$$= {}_{n-1}C_{k-1} (\theta)^{k} \cdot (1-\theta)^{n-k}$$

In [13]:
```python
# 4. (Call center) Complete the code in the iPython notebook:
def empirical_pmf(months, max_count, time_ini, time_end, verbose):
    # initializing count result
    result = np.zeros(max_count)

    # loop months and count for each months
    for month in months:
        # counts value for certain month
        counts = compute_counts(time_ini, time_end, month, verbose)

        # Feb, Apr, June, Sep, Nov has days less than 31.
        # slice counts to exact days
        if month_number[month] == 2:
            counts = counts[:28]
        elif month_number[month] == 4 or 6 or 9 or 11:
            counts = counts[:30]

        # for every day, if given day is not weekend, then add to count
        # add counts to certain result index
        for i in range(len(counts)):
            if is_weekend(month_number[month], i+1) == False:
                result[int(counts[i])] += 1

    # get final pmf
    pmf = result / sum(result)

    # return its value
    return pmf
```

```python
# import math to use factorial function
import math

def poisson_model(months, max_count,time_ini,time_end,verbose):

    # initialize two ndarrays with max_count, also initialize total calls and number of days
    pmf = np.zeros(max_count)
    result = np.zeros(max_count)
    totalCalls = 0
    numDays = 0

    # loop months and count for each months
    for month in months:
        # counts value for certain month
        counts = compute_counts(time_ini, time_end, month, verbose)

        # Feb, Apr, June, Sep, Nov has days less than 31.
        # slice counts to exact days
        if month_number[month] == 2:
            counts = counts[:28]
        elif month_number[month] == 4 or 6 or 9 or 11:
            counts = counts[:30]

        # for every day, if given day is not weekend, then count
        # add total counts to totalcalls and add 1 for each success if statement
        for i in range(len(counts)):
            if is_weekend(month_number[month], i+1) == False:
                totalCalls += counts[i]
                numDays += 1

    # lambda should be mean value, so divide total calls by number of days
    param = totalCalls / numDays

    # loop x from 0 to length of pmf, and append possion values to pmf[x]
    for x in range(len(pmf)):
        pmf[x] = (np.exp(-param) * (param ** x)) / math.factorial(x)

    # return pmf
    return pmf
```