Assignment 2: Binomial Distribution
Solutions of Problems 1,2,4 and 7

John McCuan

corrected version (Problems 4 and 7) September 18, 2023

Problem 1 (simulation; from section 2.1.1 of my notes)

(a) Use a spreadsheet program (e.g., Libre Office Calc) to generate one hundred numbers $X_1, X_2, X_3, \ldots, X_{100}$ from among the numbers in the set $\{0, 1\}$ which "appear" to be chosen randomly according to the probability measure with $\beta(\{0\}) = 1/2$.

(b) Use the spreadsheet program to count how many "heads," i.e., $X_j = 1$, you obtained.

Solution: For this problem, I decided to use Mathematica.

(a) I used the commands

```
SeedRandom[75];
expone=RandomVariate[BinomialDistribution[1, 1/2], 100] ;
Sum[expone[[j]], {j,1,100}]
```

This returned

$$\begin{array}{l}\{1,0,0,0,0,0,1,1,1,0,1,1,0,0,1,0,1,0,1,1,0,1,0,\\ \ 0,1,1,1,1,0,1,0,0,1,0,1,0,1,1,0,1,0,1,1,0,1,0,\\ \ 0,1,1,0,1,0,0,1,1,0,0,0,0,1,0,1,1,0,1,1,0,0,0,\\ \ 0,0,0,1,1,0,1,0,0,1,0,1,1,0,0,0,1,0,0,1,1,0,0,\\ \ 0,0,0,1,1,1,0,0\}\end{array}$$

46

Recall that a binomial measure with $n = 1$, which is the first parameter in `BinomialDistribution[1, 1/2]` is the Bernoulli measure. Therefore, this completes part **(a)**.

The answer to part **(b)** in this case is 46.

Problem 2 (binomial distribution; Problem 1 above)

(a) What is the probability that you did not get 50 for your answer to part **(b)** of Problem 1 above? Hint: What is the probability you did get 50?

(b) How many times would you expect to have to repeat the meaningful tasks associated with parts **(a)** and **(b)** of Problem 1 in order to see the answer "50" for (the execution of) part **(b)** at least once?

Solution: In order to get exactly 50 successes, with the first 50 entries 1 and the last 50 entries 0, the usual computation for "probability" in this case would be

$$\left(\frac{1}{2}\right)^{50}\left(\frac{1}{2}\right)^{50} = \frac{1}{2^{100}}.$$

Indeed, this is the value of the standard product measure[1]

$$\pi\left(\left\{\sum_{j=1}^{50} \mathbf{e}_j\right\}\right)$$

on $\{0,1\}^{100}$. Here we have taken $\mathbf{e}_j$ for $j = 1, 2, , \ldots, 100$ to be the standard unit basis vector as in Problem 4 below.

Just as the number of elements in $\{0,1\}^{100}$ is rather large,

$$\pi\left(\left\{\sum_{j=1}^{50} \mathbf{e}_j\right\}\right) = \frac{1}{2^{100}}$$

is rather small. However, there are other ways to get exactly 50 successes. To count those ways, we can use a combination:

$$\binom{100}{50} = 100\,891\,344\,545\,564\,193\,334\,812\,497\,256$$

which is a rather large number, more than 100 octillian.

[1]We haven't yet talked about product measures, but you might guess how they work.

If you want an explanation of how to count this number using a combination (or binomial coefficient) here is one: Say we want to choose 50 places from among the 100 entries in an element $\omega \in \{0,1\}^{100}$ to be where the 1's will appear. For the first choice, there are 100 places available. After that, there are 99 places remaining, and so on, until we have chosen the 50 places. This makes $100(99)(98)\cdots(52)(51)$ or

$$P(100,50) = \frac{100!}{(100-50)!}$$

possible ways to choose where the successes are located. We have counted them using a permutation. But we have not counted what we want. We have overcounted becuase the order in which we choose the places does not matter. If I decide succeses will happen, i.e., there will be a 1, in the first entry and the second entry of ω, it doesn't matter if I choose the first entry place first or at some other stage in the choosing. Similarly, for the second place.

As a result of this overcounting, I need to determine how many orders there are in which I can choose a particular distribution of 1's, or the places where the 50 successes will be modeled. So, assuming those 50 places are already determined, I have 50 choices (or possibilities) for the first choice, and then 49 for the second choice, and so on, until there is only one left. This means, counting using a permutation to count again, that there are

$$P(50,50) = 50!$$

ways to choose the places. Thus, dividing by this number

$$\frac{P(100,50)}{P(50,50)}$$

gives the combination used above.

The usual "probability" then is

$$\pi\left(\left\{\omega \in \{0,1\}^{100} : \#\{j \ : \ \omega_j = 1\} = 50\right\}\right) = \sum_{\omega \in F} \pi(\omega) = \#F \ \frac{1}{2^{100}}$$

where

$$F = \left\{\omega \in \{0,1\}^{100} : \#\{j \ : \ \omega_j = 1\} = 50\right\}$$

and $\#F$ is the big number $C(100, 50)$ we calculated above. That is, this "probability" is

$$\frac{12\,611\,418\,068\,195\,524\,166\,851\,562\,157}{158\,456\,325\,028\,528\,675\,187\,087\,900\,672}$$

where we have canceled a factor of $4 = 2^2$ to bring the numerator down to around 12 octillian. The denominator is still more than 158 octillian, so this number is less than $1/10$. Specifically, the probability is about 0.0795892. So if you believe in this sort of thing, you might "expect" to see exactly 50 successes, if you repeat the "random" selection of 100 elements from $\{0, 1\}$ of part **(a)** 100 times, about 8 times. Of course, this depends some on the definition of "expect," but I'll try to get into that below in my solution to part **(b)**. Put another way, you might "expect" to see exactly 50 heads once if you repeat the selection 12.5645 or 13 times. That's not so many times, so we can check it.

I executed the command

```
sum(sample(0:1, 100, replace = T))
```

in R thirteen times. Here are the results:

$$55, 52, 50, 57, 51, 51, 46, 50, 49, 44, 51, 48, 54.$$

In this case, I got exactly 50 ones on the third try and on two tries out of the thirteen.

Actually, the question asks for the "probability" that you, i.e., I, did **not** see exactly 50 ones, so this is the complementary probability:

$$\frac{145\,844\,906\,960\,333\,151\,020\,236\,338\,515}{158\,456\,325\,028\,528\,675\,187\,087\,900\,672} \doteq 0.9204011.$$

The point is that one expects to not get exactly 50 ones more than 90% of the time. And I didn't (on the first try using Mathematica), you may have. I also didn't get it on the first try using R.

In any case, I think this about finishes part **(a)**. Let's try to get into part **(b)**:

(b) Naturally, the word "expect" is a loaded word in the statement of this problem.

The interpretation of the probability 0.0795892 as indicating that if the trial described in Problem 1 is repeated some large number of times, then this will give the proportion of times that exactly 50 successes are observed is interesting. It is also slightly different, but closely related, to interpret this number as the proportion of times that 50 successes are observed out of **any** number of trials, so that the reciprocal 12.5645 is interpreted as a nominal number of trials corresponding to precisely one instance of precisely 50 successes in one of the trials. I think these would both properly be called **frequentist** interpretations.

When I think about these numbers (or when someone who doesn't believe in probability thinks about these numbers) I conclude that they really do not give me any information on what to expect. Most importantly, these numbers give no guarantee that anything in particular will be observed. Nevertheless, I did execute something that looked like 13 trials, and I saw exactly 50 ones twice. This is both consistent (with 13 trials being enough) and inconsistent (with the frequency being one). This kind of experience makes it natural and tempting (even for me) to psychologically attach some meaning to these probabilities. It is apparently tempting for many people. That is what the subject of applied probability is about after all.

Having refocused on what I really think: That the calculation of these numbers is essentially meaningless with respect to actual meaningful psychological experience (for me), I will now consider possible interpretations that I can imagine might be (and apparently are) meaningful for others.

One might say (and my son says this for example) that whenever the probability of seeing exactly 50 successes in ℓ trials is greater than 1/2, then **this** number ℓ is the number of trials one should (or at least can) use to answer this question. What this means, or seems to suggest, is that one consider the possibility (and the probability q_k) that exactly 50 successes are not observed in $k-1$ trials, but then are observed on the k-th trial for various values of $k = 1, 2, 3, \ldots$ and the smallest ℓ for which

$$\sum_{k=1}^{\ell} q_k \geq 1/2$$

is the desired number of trials.

The number $q = q_1 = 0.0795892$ we have calculated above is the probability that one sees exactly 50 successes in one trial, or on the first trial. This number is of course rather smaller than 1/2. The numbers q_k in this case correspond to a different measure (a geometric probability measure) on $\mathbb{N}$. The associated event is the execution of trials until exactly 50 successes is first observed as the outcome of the (last) trial. The (overall) outcome of this event is modeled by $k \in \mathbb{N} = \{1, 2, 3, \ldots\}$ with the outcome being the following:

something other than exactly 50 successes is observed as the (sub)outcomes of the first $k-1$ trials, and then exactly 50 successes are observed as the (sub)outcome of the k-th trial.

The measure $\gamma : \wp(\mathbb{N}) \to [0,1]$ is given by

$$q_k = \gamma(\{k\}) = (1-q)^{k-1}q.$$

As is appropriate for a probability measure we have

$$\sum_{k=1}^{\infty} q_k = \gamma(\mathbb{N}) = \sum_{k=1}^{\infty}(1-q)^{k-1}q = 1.$$

The last value is obtained because

$$\sum_{k=1}^{\infty}(1-q)^{k-1} = \sum_{j=0}^{\infty}(1-q)^j = \frac{1}{1-(1-q)} = \frac{1}{q}.$$

In fact, if $0 < \rho < 1$ and

$$s_\ell = \sum_{j=0}^{\ell} \rho^j \qquad \text{for } \ell = 2, 3, 4, \ldots, \tag{1}$$

then

$$\begin{aligned}
(1-\rho)s_\ell &= \sum_{j=0}^{\ell} \rho^j - \sum_{j=0}^{\ell} \rho^{j+1} \\
&= 1 + \sum_{j=1}^{\ell} \rho^j - \sum_{j=0}^{\ell-1} \rho^{j+1} - \rho^{\ell+1} \\
&= 1 + \sum_{j=1}^{\ell} \rho^j - \sum_{j=1}^{\ell} \rho^j - \rho^{\ell+1} \\
&= 1 - \rho^{\ell+1},
\end{aligned}$$

so that

$$s_\ell = \frac{1-\rho^{\ell+1}}{1-\rho},$$

and one may recall that the quantities s_ℓ in (1) are called **partial sums** and

$$\sum_{j=0}^{\infty} \rho^j = \lim_{\ell\to\infty} s_\ell = \lim_{\ell\to\infty} \frac{1-\rho^{\ell+1}}{1-\rho} = \frac{1}{1-\rho}.$$

In view of the foregoing calculation(s), we can find ℓ for which

$$\sum_{k=1}^{\ell} (1-q)^{k-1} q = \frac{1-(1-q)^\ell}{1-(1-q)} q = \frac{1}{2}.$$

That is,

$$(1-q)^\ell = \frac{1}{2} \qquad \text{or} \qquad \ell = \frac{\log(1/2)}{\log(1-q)} \doteq 8.35769.$$

According to this interpretation one "expects" to see exactly 50 successes after 9 trials. It is interesting that this number is different from the "frequentist interpretation" given above.

Next, I will consider what might be called **subjective confidence**. I might imagine that I am not tempted to "expect" to see exactly 50 successes in 9 trials where the cumulative probability above is a little over 1/2 simply because an alternative overall outcome is equally "likely" or almost equally likely. At which number ℓ of trials then, would I be tempted to "expect" to have seen exactly 50 successes? For example, on the face of it, were I to consider this question for only two flips per trial, then there are only four elements in the base set $\{0,1\}^2$, namely,

$$(0,0),\ (0,1),\ (1,0),\ \text{and}\ (1,1).$$

The probability of getting exactly one success is 1/2. I would certainly be disinclined to say I "expect" to see exactly one success in 1 or 2 trials, and I would never be tempted to say that. I would think, however, that at some number ℓ of trials, I would at least be tempted to say: If I flip a coin twice ℓ times, then I "expect" to have one of the trials in which there is exactly one head out of the two. To test this intuition, I executed the R command `sum(sample(0:1,2,replace = TRUE))` some number of times. The result came back one quite often. Whenever I got something other than one I would see how many trials were required to see one again. In approximately 60 trials I saw a string of four results without a one. This happened only once. There were also two or three strings of three results without a one. Obviously, the "vast majority" of the strings of trials without a "one" were rather short, and I was beginning to feel inclined or tempted to think I should expect to see exactly 1 one after 4 trials. Then the following string was returned:

$$0,\ 0,\ 2,\ 0,\ 2,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 2,\ 1. \tag{2}$$

I'll come back to this.

Here are some of the values of the (approximate) cumulative probability associated with the geometric probability discussed above.

ℓ	10	11	12	13	14	15
$1-(1-q)^\ell$	0.563668	0.598395	0.630359	0.659778	0.686856	0.711779

Table 1: Cumulative probabilities for $\ell = 10-15$ trials of "100 coin flip" experiments.

ℓ	20	25	27	28	30
$1-(1-q)^\ell$	0.809615	0.87424	0.893462	0.901941	0.916929

Table 2: Selected $\ell = 20 - 30$ trials of "100 coin flip" experiments.

ℓ	40	50	55	56	60
$1-(1-q)^\ell$	0.963753	0.984184	0.989553	0.990384	0.993099

Table 3: Selected $\ell = 40 - 60$ trials of "100 coin flip" experiments.

My initial inclination was that I would (probably) be tempted to "expect" exactly 50 ones to be seen for ℓ with probability greater than 90% or so. You can see that I calculated the value $\ell = 28$ associated with this threshold. Incidentally, the calculation of such a value is given by

$$\frac{\log(1-g)}{\log(1-q)}$$

where g is whatever probability threshold is deemed worthy of justifying "expectation," e.g., $g = 90/100$. I also calculated the value $\ell = 56$ associated with the threshold 99% (just in case).

I will also admit to being a little surprised when I saw the string of 12 trials of simulated "two coin flip" experiments with no instance of a single one as indicated in (2). Of course, I also went back and calculated the cumulative probability associated with the value $\ell = 12$ in that case, and it was over

99.9%. This confirms my inclination to avoid expectations when dealing with something I do not understand. It is perhaps worth noting that a different probabilistic/statistical point of view called Bayesianism, would probably associate "confidence" levels with the numbers in the tables above and simply offer these values as measures of "confidence" without any suggestion or interpretation of what constitutes a "high" probability. Doubletalk? I would say most likely, "yes."

Finally, concerning the outcome one "expects," we have noted before, that the average value of a real valued (payoff) function over a probability measure space, i.e., the integral, is also called the "expectation" of the value of that function. We can ask the question: Is there a probability measure for which one of these values is the expectation? The answer is "yes." The expectation of the identity function $\mathrm{id}_{\mathbb{R}}$ with respect to the geometric probability measure is

$$\int_{\mathbb{R}} \mathrm{id} = \int_{\xi \in \mathbb{R}} \xi = \sum_{j \in \mathbb{N}} j\, \gamma(\{j\}) = \frac{1}{q} \doteq 12.5645.$$

So in the technical sense of the language associated with probability, the "expectation" is about 13 trials. I'm not sure this means one should "expect" to see exactly 50 ones after 13 trials, but I guess there is a probability of about 66% for that outcome. To calculate the integral, we can use geometric series and the fact that series of non-negative terms can be rearranged in any manner.

Specifically,

$$
\begin{aligned}
\sum_{j=1}^{\infty} j(1-q)^{j-1}q &= q\sum_{j=1}^{\infty}\sum_{\ell=1}^{j}(1-q)^{j-1} \\
&= q\sum_{\ell=1}^{\infty}\sum_{j=\ell}^{\infty}(1-q)^{j-1} \\
&= q\sum_{\ell=1}^{\infty}(1-q)^{\ell-1}\sum_{j=\ell}^{\infty}(1-q)^{j-\ell} \\
&= q\sum_{\ell=1}^{\infty}(1-q)^{\ell-1}\sum_{k=0}^{\infty}(1-q)^{k} \\
&= q\sum_{\ell=1}^{\infty}(1-q)^{\ell-1}\frac{1}{1-(1-q)} \\
&= \sum_{\ell=1}^{\infty}(1-q)^{\ell-1} \\
&= \frac{1}{1-(1-q)} \\
&= \frac{1}{q}.
\end{aligned}
$$

Problem 3 (Problem 2 above) If you calculated a probability in order to answer part **(b)** of Problem 2 above, what was the base model set S for your calculation? What is interesting about this set?

Solution: As a matter of fact, I calculated all kinds of probabilities, but the basic intention of this problem was to see if you (the student) invoked the gometric probability measure associated with the probability q. As mentioned above, the base model set for that probability in my case was $\mathbb{N}$.

The interesting thing about this set is that it is not a set with finitely many elements. (At least that is the interesting thing I imagined you might appreciate if you were thinking about the problem as I was above. Most of the model sets considered here are measure spaces with finitely many elements, but this one has infinitely many elements corresponding to any arbitrary number of "trials.")

Problem 4 (simulation; from section 2.1.1 of my notes)

(a) Use a spreadsheet program (e.g., Libre Office Calc) to generate one hundred numbers $X_1, X_2, X_3, \dots, X_{100}$ from among the numbers in the set $\{0, 1\}$ which "appear" to be chosen randomly according to the probability measure with $\beta(\{0\}) = 1/3$.

(b) Define a base set $S \subset \mathbb{R}^{100}$ to model the possible outcomes of the event that took place when you executed part **(a)** above.

(c) Define a probability measure $\pi : \wp(S) \to [0, 1]$ on S with

$$\pi\left(\left\{ \sum_{j \in A} \mathbf{e}_j \right\} \right) = \left(\frac{2}{3}\right)^{\#A} \left(\frac{1}{3}\right)^{100-\#A}$$

for every $A \subset \{1, 2, 3, \dots, 100\}$, and define a function $y : S \to \mathbb{N}$ by

$$y\left(\sum_{j \in A} \mathbf{e}_j \right) = \sum_{j \in A} j.$$

In these expressions $\mathbf{e}_j$ denotes the standard unit basis vector in $\mathbb{R}^{100}$. Find the value of the integral of the function y over S with respect to the measure π, and explain the meaning of this number.

Correction: The definition of the function $y : S \to \mathbb{N}$ above is supposed to be

$$y\left(\sum_{j \in A} \mathbf{e}_j \right) = \sum_{j \in A} 1.$$

The problem with the definition above makes sense, but it is quite difficult at least to do by hand. Maybe I intended for you to do it with mathematical software like Mathematica or a spreadsheet program or something, but I think I just forgot to change the "j" to a "1."

Addition: For those not familiar with the standard unit basis vectors, here is a little additonal explantion. The vector $\mathbf{e}_1 = (1, 0, 0, 0, \dots)$ with 100 entries and the first one is 1. Similarly, $\mathbf{e}_2 = (0, 1, 0, 0. \dots)$ with a 1 in the second component/entry. These can be added to get $\mathbf{e}_1 + \mathbf{e}_2 = (1, 1, 0, 0, 0, \dots) \in \{0, 1\}^{100}$.

Solution: For this problem, I decided to use Mathematica.

(a) The command

```
SeedRandom[123]; RandomVariate[BinomialDistribution[1, 2/3], 100]
```

returned

$$\{1,1,1,1,0,1,0,1,1,1,0,1,0,0,1,0,0,1,0,0,1,1,1,\\ 0,0,0,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,\\ 0,1,1,0,1,0,1,1,1,1,1,1,1,1,0,0,1,1,1,1,1,1,1,\\ 1,0,1,1,0,0,1,1,0,0,1,1,1,0,1,1,1,1,1,1,0,0,1,\\ 1,1,0,0,0,1,0,0\}$$

Recall that a binomial measure with $n = 1$, which is the first parameter in `BinomialDistribution[1, 2/3]` is the Bernoulli measure. Therefore, this completes part **(a)**.

(b) $S = \{0, 1\}^{100}$.

(c) Notice that π is precisely the binomial measure, and y is the function that induces the binomial distribution. Thus, we are being asked to calculate the average value of y with respect to the binomial measure:

$$\int_S y = \sum_{\omega \in S} y(\omega)\ \pi(\{\omega\}).$$

Of course, there are a lot (2^{100}) elements in S, so we need to do something somewhat clever here (to calculate this by hand).

One thing we can do is group all the values together which are determined by a particular number $\#A$ giving the number of times 1 appears in a particular value of ω. If we do this, we see

$$\int_S y = \sum_{k=0}^{\infty} k \, \binom{100}{k} \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right)^{100-k} \tag{3}$$

where we have taken $k = \#A$.

Recalling the definition of the induced measure α_y, this is the same as

$$\int_S y = \sum_{\xi\in\mathbb{R}} \xi\ \alpha_y(\{\xi\}) = \int_{\xi\in\mathbb{R}} \xi$$

where the integral on the right is taken with respect to the induced measure. In this way we see the number we are asked to compute is also the average (or the **mean**) of the binomial distribution for $n = 100$, but none of this helps us much to find the actual value.

One way to get at the actual value is to consider y as a sum of other functions $y_\ell : S \to \mathbb{R}$ for $\ell = 1, 2, 3, \ldots, 100$. Namely, we can take

$$y_\ell(\omega) = \#(A \cap \{\ell\}) \qquad \text{where} \qquad \omega = \sum_{j\in A} \mathbf{e}_j.$$

That is,

$$y_\ell(\omega) = \begin{cases} 0, & \omega_\ell = 0 \\ 1, & \omega_\ell = 1. \end{cases}$$

On the one hand, it is clear from this description that

$$y = \sum_{\ell=1}^{100} y_\ell \qquad \text{and} \qquad \int_S y = \sum_{\ell=1}^{100} \int_S y_\ell.$$

On the other hand, the functions y_ℓ for $\ell = 1, 2, \ldots, 100$ may just be pretty easy to integrate. Let's see.

$$\begin{aligned} \int_S y_1 &= \sum_{\omega\in S, \omega_1=1} y_\ell(\omega)\ \pi(\{\omega\}) \\ &= \sum_{\omega\in S, \omega_1=1} \pi(\{\omega\}). \end{aligned} \tag{4}$$

The value of $\pi(\{\omega\})$ will, of course, depend on the entries $\omega_2, \ldots, \omega_{100}$ in ω, but we at least know there is one factor of $2/3$ for the first entry. Therefore, we see

$$\pi(\{\omega\}) = \frac{2}{3}\left(\frac{2}{3}\right)^{\#\{j\ :\ \eta_j=1\}} \left(\frac{1}{3}\right)^{99-\#\{j\ :\ \eta_j=1\}}$$

where $\eta \in \{0,1\}^{99}$ and $\omega = (1, \eta)$. Letting β denote the binomial measure for $n = 99$, this becomes also

$$\pi(\{\omega\}) = \frac{2}{3}\beta(\{\eta\}).$$

With this in mind, we may continue from (4) to conclude

$$\begin{aligned}\int_S y_1 &= \sum_{\omega \in S, \omega_1 = 1} \frac{2}{3}\beta(\{\eta\}) \\ &= \frac{2}{3} \sum_{\eta \in S_{99}} \beta(\{\eta\}) \\ &= \frac{2}{3}\beta(S_{99})\end{aligned}$$

where $S_{99} = \{0,1\}^{99}$. But β is a probability measure on $S_{99} = \{0,1\}^{99}$, so

$$\int_S y_1 = \frac{2}{3}.$$

Essentially the same computation applies for each $\ell = 2, 3, \ldots, 100$ so that

$$\int_S y_\ell = \frac{2}{3} \qquad \text{for } \ell = 1, 2, 3, \ldots, 100$$

as well. Finally, then

$$\begin{aligned}\int_S y &= \sum_{\ell=1}^{100} \int_S y_\ell \\ &= \sum_{\ell=1}^{100} \frac{2}{3} \\ &= \frac{200}{3} = 66 + \frac{2}{3}.\end{aligned}$$

I will repeat this calculation in a more general case below and calculate the integral for the more complicated function $z : S_n \to \mathbb{R}$ by

$$z(\omega) = z\left(\sum_{j \in A} \mathbf{e}_j\right) = \sum_{j \in A} j$$

given accidentally in the original statement of the problem.

As for meaning, the integral we have calculated is the expectation or average one might expect for the sum of the 100 entries when the event of part **(a)** takes place. In this particular, case, the outcome of the pseudo-random sample

$$\begin{aligned}\omega_* =&(1,1,1,1,0,1,0,1,1,1,0,1,0,0,1,0,0,1,0,0,1,1,1,\\&0,0,0,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,\\&0,1,1,0,1,0,1,1,1,1,1,1,1,1,0,0,1,1,1,1,1,1,1,\\&1,0,1,1,0,0,1,1,0,0,1,1,1,0,1,1,1,1,1,1,0,0,1,\\&1,1,0,0,0,1,0,0)\end{aligned}$$

obtained using Mathematica for part **(a)** has

$$y(\omega_*) = 68.$$

Exercise: Let $S = \{0, 1\}$ so that $S^n = \{0, 1\}^n$ for some $n \in \mathbb{N}$. Generalize the computation above for the integral of $x : S^n \to \mathbb{R}$ by

$$x(\omega) = \sum_{\ell=1}^{n} \omega_\ell$$

with respect to the binomial measure β_n and any probability p with $0 < p < 1$. Show

$$\int_{S^n} x = \int_{\xi\in\mathbb{R}} \xi$$

where the integral on the right is with respect to the measure α_n induced by x.

Solution: To see the last integral identity note that

$$
\begin{aligned}
\int_{S^n} x &= \sum_{\omega\in S^n} x(\omega)\ \beta_n(\{\omega\}) \\
&= \sum_{\omega\in S^n} x(\omega)\ p^{\#\{j\ :\ \omega_j=1\}}(1-p)^{n-\#\{j\ :\ \omega_j=1\}} \\
&= \sum_{\omega\in S^n} \#\{j\ :\ \omega_j=1\} p^{\#\{j\ :\ \omega_j=1\}}(1-p)^{n-\#\{j\ :\ \omega_j=1\}} \\
&= \sum_{k=0}^{n} \binom{n}{k} kp^k(1-p)^{n-k} && (5) \\
&= \sum_{k=0}^{n} k\ \alpha_n(\{k\}) && (6) \\
&= \sum_{\xi\in\mathbb{R}} \xi\ \alpha_n(\{\xi\}) \\
&= \int_{\xi\in\mathbb{R}} \xi.
\end{aligned}
$$

To get (5) we take $k = \#\{j\ :\ \omega_j = 1\}$ and/or group points $\omega \in S^n$ together for which $\#\{j\ :\ \omega_j = 1\} = k$. One must ask how many points ω are there in S^n with $\#\{j\ :\ \omega_j = 1\} = k$, that is, one must compute

$$\#\big\{\omega \in S^n :\ \#\{j\ :\ \omega_j = 1\} = k\big\} = \binom{n}{k}.$$

The expression (6) uses the definition of the binomial induced measure $\alpha = \alpha_n$.

Here, I will put the calculation all together.

$$
\begin{aligned}
\int_{S^n} x &= \int_{\omega \in S^n} \sum_{\ell=1}^{n} \omega_\ell \\
&= \sum_{\ell=1}^{n} \int_{\omega \in s^n} \omega_\ell \\
&= \sum_{\ell=1}^{n} \int_{s^n} x_\ell \quad \text{where } x_\ell : S^n \to \mathbb{R} \text{ by } x_\ell(\omega) = \omega_\ell \\
&= \sum_{\ell=1}^{n} \sum_{\omega \in S^n} \omega_\ell \beta_n(\{\omega\}) \\
&= \sum_{\ell=1}^{n} \sum_{\omega \in S^n, \omega_\ell = 1} \beta_n(\{\omega\}) \\
&= \sum_{\ell=1}^{n} \sum_{\omega \in S^n, \omega_\ell = 1} (p)^{\#\{j \,:\, \omega_j = 1\}} (1-p)^{n - \#\{j \,:\, \omega_j = 1\}} \\
&= \sum_{\ell=1}^{n} \sum_{\omega \in S^n, \omega_\ell = 1} (p)^{1 + \#\{j \neq \ell \,:\, \omega_j = 1\}} (1-p)^{n - 1 - \#\{j \neq \ell \,:\, \omega_j = 1\}} \\
&= p \sum_{\ell=1}^{n} \sum_{\omega \in S^n, \omega_\ell = 1} (p)^{\#\{j \neq \ell \,:\, \omega_j = 1\}} (1-p)^{n - 1 - \#\{j \neq \ell \,:\, \omega_j = 1\}} \\
&= p \sum_{\ell=1}^{n} \sum_{\omega \in S^{n-1}} (p)^{1 + \#\{j \,:\, \omega_j = 1\}} (1-p)^{n - 1 - \#\{j \,:\, \omega_j = 1\}} \\
&= p \sum_{\ell=1}^{n} \sum_{\omega \in S^{n-1}} \beta(\{\omega\}) \\
&= p \sum_{\ell=1}^{n} \beta_{n-1}(S^{n-1}) \\
&= p \sum_{\ell=1}^{n} 1 \\
&= pn.
\end{aligned}
$$

Alternative calculation (Jeremy Mahoney):

Jeremy obtained the form (5) and proceeded to calculate the value directly. His derivation was based on two combinatorics identities. The first gives an alternative form for the factor

$$k \binom{n}{k}$$

in (5) namely

$$k \binom{n}{k} = n \binom{n-1}{k-1}. \tag{7}$$

This allows factoring n from the sum. The second identity is fairly well-known and should appear in some form in my chapter/lecture 3 notes:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}. \tag{8}$$

This is called **Pascal's identity**.[2] I think I verify Pascal's identity in my chapter/lecture 3 notes. I don't think it is difficult: Just expand the summed combinations on the right (8) get a common denominator, and add. One thing it turns out is important to notice is that one needs $k < n$ in order to properly state and use Pascal's identity. I wonder if I stated this properly in my notes.

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \qquad (k < n). \tag{9}$$

This point came up as a small glitch in Jeremy's presentation, and I should be able to tidy that up below.

Uma Anand gave a heuristic explanation/dervation of Pascal's identity (9) which was quite clear and compelling—again using terminology which is somewhat new to me, but relatively easy to figure out or look up. If you know the terminology of "bit strings" and are familiar with counting them, then I think Uma's explanation of Pascal's identity can be summed up in one sentence:

> Adding the number of bit strings of length n with 0 as first bit to the number of bit strings with 1 as first bit, you get Pascal's identity.

[2] This is new terminology to me, but that is what Jeremy called it, and it is a really good name as it illustrates the construction of Pascal's triangle—which I do know about. I guess people who know some combinatorics know these things, and I am happy to learn them.

For those with less famiarity with this sort of thing, which is maybe only me, I offer some additional explanation:

Uma's basic idea is to interpret the combination

$$\binom{n}{k}$$

generally as the number of "bit strings" of length n, i.e., elements $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ of $\{0, 1\}^n$, containing k bit[3] locations in which one finds the value 1. We can perhaps call this the **general principle of counting bit strings**. It works whenever $k \leq n$.

Then Uma suggests counting the same collection of bit strings of length n having exactly k locations containing a 1 in a different way:

If $k < n$ there are two choices for the first bit, i.e., ω_1 in $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$. If $k = n$, there are not two choices for the first bit, and the entire counting problem sort of collapses: There is only one bit string of length n with n ones.

Returning to the case $k < n$, if you take the first bit to be 1, then there are $k-1$ remaining 1's for which locations must be found among the remaining $n-1$ bit locations. The general principle of counting bit strings applies to give

$$\binom{n-1}{k-1}$$

such bit strings. Of course, one can still do this part when $k = n$, but it doesn't give you anything particularly interesting.

If $k < n$ however, it is possible to take the first bit to be 0, and it makes sense to count the number of bit strings of length $n-1$ having a 1 in k bit locations. Again the general principle of counting bit strings applies to give

$$\binom{n-1}{k}$$

such bit strings. Thus, we can repeat the summary:

> Adding the number of bit strings of length n with 0 as first bit to the number of bit strings with 1 as first bit, you get Pascal's identity.

[3]More properly, a "bit" is a "portmanteau" of "binary digit," that is a "0 or a 1" in this case one of n ordered zeros and ones in an n-tuple. I'm sure you know all this, except perhaps the word "portmanteau" which is from linguistics.

Here is Jeremy's derivation of his first identity (7):

$$\begin{aligned} k \binom{n}{k} &= k \, \frac{n!}{k!(n-k)!} \\ &= n \, \frac{(n-1)!}{(k-1)!(n-k)!} \\ &= n \, \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} \\ &= n \binom{n-1}{k-1}. \end{aligned}$$

With these tools in hand, we attack the expression (5):

$$\begin{aligned} \int_{S^n} x &= \sum_{k=0}^{n} k \, \binom{n}{k} \, p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} n \, \binom{n-1}{k-1} \, p^k (1-p)^{n-k} \qquad \text{Jeremy's first identity} \\ &= n \sum_{k=0}^{n} \binom{n-1}{k-1} \, p^k (1-p)^{n-k} \\ &= n \left(\sum_{k=0}^{n-1} \binom{n-1}{k-1} \, p^k (1-p)^{n-k} \qquad + \qquad p^n \right) \\ &= n \left(\sum_{k=0}^{n-1} \left[\binom{n}{k} - \binom{n-1}{k} \right] \, p^k (1-p)^{n-k} + p^n \right) \qquad \text{Pascal's identity} \\ &= n \sum_{k=0}^{n-1} \binom{n}{k} \, p^k (1-p)^{n-k} - n \sum_{k=0}^{n-1} \binom{n-1}{k} \, p^k (1-p)^{n-k} + np^n \\ &= n \sum_{k=0}^{n} \binom{n}{k} \, p^k (1-p)^{n-k} - n \sum_{k=0}^{n-1} \binom{n-1}{k} \, p^k (1-p)^{n-k}. \end{aligned}$$

The summation

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

is precisely the binomial expansion formula for $[a+b]^n$ with $a = p$ and $b = 1-p$.

Therefore, this sum is $[p + (1-p)]^n = 1^n = 1$. The other summation

$$\sum_{k=0}^{n-1} \left(\begin{array}{c} n-1 \\ k \end{array} \right) p^k (1-p)^{n-k}$$

is also nearly a simple binomial expansion:

$$\begin{aligned} \sum_{k=0}^{n-1} \left(\begin{array}{c} n-1 \\ k \end{array} \right) p^k (1-p)^{n-k} &= (1-p) \sum_{k=0}^{n-1} \left(\begin{array}{c} n-1 \\ k \end{array} \right) p^k (1-p)^{n-1-k} \\ &= (1-p)[p + (1-p)]^{n-1} \\ &= (1-p). \end{aligned}$$

With these observations, we can complete the calcuation:

$$\sum_{k=0}^{n} k \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} = n[1 - (1-p)] = np.$$

Comment: You can see that I've computed a somewhat more complicated integral over S^n below, and I don't think the first step of expressing this more complicated integral as an integral over $\mathbb{R}$ with respect to the induced measure α can be completed. At least I do not see a way to do it, and I see a reason why it is not a very natural thing to do. If I'm correct about this, I guess it also means that this kind of direct combinatoric approach will either not quite work or not be quite as straightforward. But maybe I'm wrong about that.

Finally, let's use the same techniques to integrate the more complicated function $z : S^n \to \mathbb{R}$ given by

$$z(\omega) = z\left(\sum_{j \in A} \mathbf{e}_j \right) = \sum_{j \in A} j.$$

Here we are taking the general case with $A \subset \{1, 2, \ldots, n\}$ specifying a point in S^n by

$$\omega = \sum_{j \in A} \mathbf{e}_j.$$

We first note that

$$z(\omega) = \sum_{\ell=1}^{n} \ell \omega_\ell = \sum_{\ell=1}^{n} z_\ell$$

where $z_\ell : S^n \to \mathbb{R}$ by $z_\ell(\omega) = \ell\omega_\ell$ for $\ell = 1, 2, \ldots, n$. Thus, we can still use the linearity of the integral. Note also that the computation above gives

$$\sum_{\omega \in S^n, \omega_\ell = 1} \beta_n(\{\omega\}) = p \tag{10}$$

in general. Thus, we may compute as follows:

$$\begin{aligned}
\int_{S^n} z &= \int_{\omega \in S^n} \sum_{\ell=1}^{n} \ell\omega_\ell \\
&= \sum_{\ell=1}^{n} \int_{\omega \in s^n} \ell\omega_\ell \\
&= \sum_{\ell=1}^{n} \int_{s^n} z_\ell \quad \text{where } z_\ell : S^n \to \mathbb{R} \text{ by } z_\ell(\omega) = \ell\omega_\ell \\
&= \sum_{\ell=1}^{n} \sum_{\omega \in S^n} \ell\omega_\ell \beta_n(\{\omega\}) \\
&= \sum_{\ell=1}^{n} \ell \sum_{\omega \in S^n, \omega_\ell = 1} \beta_n(\{\omega\}) \\
&= \sum_{\ell=1}^{n} \ell p \\
&= p \sum_{\ell=1}^{n} \ell \\
&= p\ \frac{n(n+1)}{2}.
\end{aligned}$$

It might also be nice to realize this more complicated integral as the integral of an appropriate function over $\mathbb{R}$ with respect to the induced measure $\alpha = \alpha_n$ but I do not see an obvious way to do that.

$$
\begin{aligned}
\int_{S^n} z &= \sum_{\omega \in S^n} z(\omega)\ \beta_n(\{\omega\}) \\
&= \sum_{\omega \in S^n} z(\omega)\ p^{\#\{j\ :\ \omega_j = 1\}}(1-p)^{n-\#\{j\ :\ \omega_j=1\}} \\
&= \sum_{\omega \in S^n} \sum_{\ell=1}^{n} \ell \omega_\ell p^{\#\{j\ :\ \omega_j = 1\}}(1-p)^{n-\#\{j\ :\ \omega_j=1\}} \\
&= \sum_{\ell=1}^{n} \ell \sum_{\omega \in S^n} \omega_\ell p^{\#\{j\ :\ \omega_j = 1\}}(1-p)^{n-\#\{j\ :\ \omega_j=1\}} \\
&= \sum_{\ell=1}^{n} \ell \sum_{k=0}^{n} \sum_{\omega \in Q_k} \omega_\ell p^k (1-p)^{n-k} \\
&= \sum_{\ell=1}^{n} \ell \sum_{k=1}^{n} \sum_{\omega \in Q_k, \omega_\ell = 1} p^k (1-p)^{n-k} \\
&= \sum_{\ell=1}^{n} \ell \sum_{k=1}^{n} \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&= \sum_{\ell=1}^{n} \ell p \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-1-(k-1)} \\
&= \sum_{\ell=1}^{n} \ell p \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k} (1-p)^{n-1-k} \\
&= \sum_{\ell=1}^{n} \ell p \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k} (1-p)^{n-1-k} \\
&= p \sum_{\ell=1}^{n} \ell\ \sum_{k=0}^{n-1} \alpha_{n-1}(\{k\}) \\
&= p \sum_{\ell=1}^{n} \ell\ \alpha_{n-1}(\mathbb{R}).
\end{aligned}
$$

$$
Q_k = \big\{\omega \in S^n : \#\{j\ :\ \omega_j = 1\} = k\big\}.
$$

Problem 5 (sections 2.1.2 and 2.1.3 in my notes) Let $\beta : \mathcal{P}(\mathbb{R}) \to [0,1]$ denote the Bernoulli measure. Let $x : \mathbb{R} \to \mathbb{R}$ by $x(t) = t + 5$ be a renaming bijection.

(a) Find the measure induced by the renaming x.

(b) Find the PMF of the induced measure obtained in part **(a)** above. How is this PMF different from the PMF of the Bernoulli measure?

Problem 6 (section 2.1.3 and 2.1.4 in my notes, especially Exercise 2.1.8; also see the discussion of Orloff and Booth concerning intransitive dice) Let

$$S = \{\text{one}, \text{two}, \text{three}, \text{four}, \text{five}, \text{six}\}.$$

Consider the uniform probability measure π on S. Now, say you have three (standard six sided) dice—well, not completely standard: One is red and has five sides marked "three," i.e., with three dots, and one side marked "six." A second die is green and has one side marked "one" and five sides marked "four." The third die is blue with three sides marked "two" and three sides marked "five."

(a) Define three real valued functions $r : S \to \mathbb{R}$, $g : S \to \mathbb{R}$ and $b : S \to \mathbb{R}$ appropriate for modeling the possible outcomes of rolling the red, green, and blue die respectively.

(b) Find the induced measures α_r, α_g and α_b associated with each of the functions from part **(a)** above and plot the associated PMF.

(c) Find the integrals of r, g and s with respect to π.

Problem 7 (Problem 6 above) Consider the red, green and blue dice of Problem 6 above. Let

$$C = S \times S = \{(\omega_1, \omega_2) : \omega_j \in S\}$$

be a set used to model the outcome of a competition between a pair of the three dice from Problem 6.

(a) Describe the uniform measure on C.

(b) What function $h : C \to \mathbb{R}^2$ is appropriate to model the outcome of a competition between the red die and the blue die?

(c) Using your function from part **(b)** find the measure α_w induced on C by $w = \delta \circ h : C \to \mathbb{R}$ where $\delta(\xi_1, \xi_2) = \xi_2 - \xi_1$.

Correction: An induced measure measures the subsets in the codomain of the inducing function. In this case, the induced measure should be a measure on $\mathbb{R}$, so if I want this to make sense, I should write

> Using your function from part **(b)** find the measure α_w induced on $\mathbb{R}$ by $w = \delta \circ h : C \to \mathbb{R}$ where $\delta(\xi_1, \xi_2) = \xi_2 - \xi_1$.

Having made this correction, let me mention something: The induced measure involves preimages under w back in the set C, and I think it is important to see those preimages back in C and the corresponding values of the base (uniform) measure. That's what I had in mind. Perhaps I should add an additional part making this explicit, but I'll leave it for now.

(d) Using your induced measure $\alpha_w : \wp(C) \to [0,1]$ what is the value and meaning of

$$\alpha_w \{ \ \{(\omega_1, \omega_2) : w(\omega_1, \omega_2) > 0\} \ )? \tag{11}$$

Correction: This part[4] should read

> Using your induced measure $\alpha_w : \wp(\mathbb{R}) \to [0,1]$ what is the value and meaning of
>
> $$\alpha_w(\{\xi \in \mathbb{R} : \xi > 0\})?$$

You'll note that there is also an incorrect grouping symbol in (11); the first "curly bracket" should just be round evaluation parentheses. Incidentally, the set

$$\{(\omega_1, \omega_2) \in C : w(\omega_1, \omega_2) > 0\}$$

appearing in (11) is essentially the set I'd like to make sure you "see."

(e) Repeat parts **(b)** and **(c)** for competitions between the remaining pairs of dice. What interesting thing do you find?

Solution:

(a) There are 36 elements in the competition set $C = S \times S$, so the uniform (probability) measure $\pi : \wp(C) \to [0,1]$ has values given by

$$\pi(\{(\omega_1, \omega_2)\}) = \frac{1}{36} \qquad \text{and} \qquad \pi(A) = \frac{\#A}{36}.$$

(b) There might be a couple possible answers to this question in general, but the way I have it phrased with the codomain $\mathbb{R}^2$ for h, the answer I had in mind is $h : C \to \mathbb{R}^2$ by

$$h(\omega_1, \omega_2) = (r(\omega_1), g(\omega_2)).$$

If you don't get this answer (the corrected) part **(c)** will be tough.

[4] if I want it to be correct and make sense

(c) Notice first that according to my choice of h in part **(b)**, I get

$$w(\omega_1, \omega_2) = g(\omega_2) - r(\omega_1).$$

That is, w is positive if the green die shows a higher value and is negative if the red die shows a higher value. If you did part **(a)** of Problem 6 the way I did it, the values of the red die are in the codomain of the function r and the values of the green die are in the codomain of the function g so that for example if $r(\text{six}) = 6$, then

$$g(\omega_2) - r(\text{six}) = g(\omega_2) - 6$$

can take precisely two values: -5 or -2. All the possible image values of w together are given by the set $\{-5, -2, 1\}$.

If we want to understand the induced measure we should consider $w^{-1}(\{-5\})$ and $w^{-1}(\{-2\})$ corresponding to situations in which

> "the red die shows a higher value than the green die,"

and $w^{-1}(\{1\})$ corresponding to situations in which "the green die shows a higher value than the red die." The way I did it the following array illustrates the set(s) in C I was hoping you would find to be of primary interest:

$$\begin{array}{rl} g=4 & \left\{\begin{array}{ccccc} (1,6) & (2,6) & \cdots & (5,6) & (6,6) \\ \vdots & & & \vdots & \vdots \\ (1,3) & (3,3) & \cdots & (5,1) & (6,3) \\ (1,2) & (2,2) & \cdots & (5,1) & (6,2) \end{array}\right. \\ g=1 & \left\{ \underbrace{(1,1) \quad (2,1) \quad \cdots \quad (5,1)}_{r=3} \quad \underbrace{(6,1)}_{r=6} \right. \end{array}$$

Note that in the array, I've used numbers/numerals to represent the number words "one," "two," and so forth. From the array, I hope you can see

$$w^{-1}(\{-5\}) = \{(\text{six}, \text{one})\}.$$

This is a set with measure $1/36$ back in C. Thus, $\alpha_w(\{-5\}) = 1/36$. Similarly,

$$w^{-1}(\{-2\}) = \{(\text{one}, \text{one}), (\text{two}, \text{one}), \ldots, (\text{five}, \text{one}), (\text{six}, \text{two}), (\text{six}, \text{two}), \ldots, (\text{six}, \text{two})\}.$$

Consequently,

$$\alpha_w(\{-2\}) = \frac{10}{36} = \frac{5}{18}.$$

Finally, the large set where $w = 1$ gives

$$\alpha_w(\{1\}) = \frac{25}{36}.$$

Any set A with $A \subset \mathbb{R}\backslash\{-5, -2, 1\}$ has $\alpha_w(A) = 0$. We have specified the induced measure completely. In particular, the PMF of this measure is given by

$$M(\xi) = \begin{cases} 0, & \xi \neq -5, -2, 1 \\ 1/36, & \xi = -5 \\ 5/18, & \xi = -2 \\ 25/36, & \xi = 1. \end{cases}$$

(d) According to the discussion of part **(c)**, we have

$$\alpha_w(\{\xi \in \mathbb{R} : \xi > 0\}) = \alpha_w(\{1\}) = \frac{25}{36}.$$

That is the set corresponding to "the greeen die shows a value greater than the red die" has measure $25/36 >> 18/36 = 1/2$.

(e) Maybe I should have said "Repeat parts **(b)**, **(c)** and **(d)** for competitions between the remaining pairs of dice." In any case, let me change notation slightly and indicate what I had in mind. Now we have several competitions. We can model each competititon with $S \times S$, but the physical meaning (or modeling) in each case will be quite different. Accordingly, I suggest a subscript "1" be used in reference to the competition of "red against green" considered above, so that the function induced measure is denoted by α_1 instead of α_w.

In considering "red against blue," the values of the blue die are 2 and 5 with, say $b(j) = 2$ for $j = 1, 2, 3$ and $b(j) = 5$ for $j = 4, 5, 6$, again using numbers for number names. We find therefore a measure α_2 induced by $w_2(\omega_1, \omega_2) = b(\omega_2) - r(\omega_1)$ with support $\{-4, -1, 2\} \subset \mathbb{R}$ since the values taken by w_2 are -4, -1, and -2. Computing as above (if I've done it correctly) I see

$$\alpha_2(\{-4\}) = \frac{3}{36} = \frac{1}{12}, \quad \alpha_2(\{-1\}) = \frac{18}{36} = \frac{1}{2}, \quad \text{and} \quad \alpha_2(\{2\}) = \frac{15}{36} = \frac{5}{12}.$$

Here we can say

$$\alpha_2(\{\xi \in \mathbb{R} : \xi > 0\}) = \alpha_2(\{2\}) = \frac{5}{12} < \frac{1}{2},$$

but perhaps more interesting is the observation that

$$\alpha_2(\{\xi \in \mathbb{R} : \xi < 0\}) = \frac{7}{12} > \frac{1}{2}.$$

This means the set corresponding to "the red die shows a value greater than the blue die" is $7/12 > 1/2$.

Combining this with our conclusion from before we have

> The set corresponding to "green greater than red" has measure greater than $1/2$, and the set corresponding to "red greater than blue" has measure greater than $1/2$.

One might be tempted to conclude from this that the set corresponding to "green greater than blue" also has measure greater than $1/2$, but one should keep in mind that whatever the measure of that set might be, it is the result of a third competition where the underlying model sets are modeling different dice, and the value of a third and different induced measure is involved.

So we should check it. Let's say "green competes against blue." The values for green are 4 and 1 while the values for blue are 2 and 5. Taking w_3 with

$$w_3(\omega_1, \omega_2) = b(\omega_2) - g(\omega_1)$$

we find w_3 takes the values -2, 1, and 4. Also,

$$\alpha_3(\{-2\}) = \frac{15}{36} = \frac{5}{12}, \quad \alpha_3(\{1\}) = \frac{18}{36} = \frac{1}{2}, \quad \text{and} \quad \alpha_3(\{4\}) = \frac{3}{36} = \frac{1}{12}.$$

Consequently, the set

$$w_3^{-1}(\{\xi \in \mathbb{R} : \xi > 0\}) = \{(\omega_1, \omega_2) \in S \times S : b(\omega_2) - g(\omega_1) > 0\}$$

corresponding to "blue is greater than green" has measure $21/36 = 7/12 > 1/2$. Alternatively,

$$\alpha_3(\{\xi \in \mathbb{R} : \xi > 0\}) = \frac{7}{12} > \frac{1}{2}.$$

We conclude then, possibly contrary to our intuition, that the set corresponding to "green is greater than blue" does **not** have measure greater than $1/2$.

Problem 8 (From sections 2.4.1 and 2.4.2 in my notes)

(a) Plot the PMF of the binomial distribution when $p = 1$ and $n = 1, 2, 3, \ldots$.

(b) Plot the CMF of the binomial distribution when $p = 1/2$ and $n = 2, 3, 4$.

Problem 9 (binomial distribution) Let β denote the binomial measure and α_x the binomial induced measure.

(a) Calculate

$$\lim_{p \searrow 0} \alpha_x(\{0\}) \qquad \text{and} \qquad \lim_{p \searrow 0} \alpha_x(\{1\}).$$

(b) Calculate

$$\lim_{p \nearrow 1} \alpha_x(\{0\}) \qquad \text{and} \qquad \lim_{p \nearrow 1} \alpha_x(\{1\}).$$

(c) Find the unique value p_0 of p for which $\alpha_x(\{0\}) = \alpha_x(\{1\})$.

(d) What is the relation between $\alpha_x(\{0\})$ and $\alpha_x(\{1\})$ when $p < p_0$?

(d) What is the relation between $\alpha_x(\{0\})$ and $\alpha_x(\{1\})$ when $p > p_0$?

Problem 10 (binomial distribution; Problem 9 above) Let β denote the binomial measure and α_x the binomial induced measure.

(a) Find the unique value p_* of p for which

$$\alpha_x(\{0\}) = \sum_{j=1}^{n} \alpha_x(\{j\}). \tag{12}$$

(b) If p is the probability of "success" in an event modeled by the Bernoulli base set $\{h, t\}$ and $\alpha_x(\{\xi\})$ is the probability of having ξ "successes" out of n Bernoulli trials, describe the probabilistic interpretation of the inequality $p < p_*$.