TOPIC 12: RANDOM VARIABLES AND THEIR DISTRIBUTIONS

In the last section we compared the length of the longest run in the data for various players to our expectations for the longest run in data generated in independent Bernoulli Trials with a given value of p. The idea was that if the longest run in the given set of data was "too long" or "too short" we could conclude that the data was not generated as a result of independent Bernoulli trials with the given value of p. Of course, we could not conclude beyond a shadow of doubt that this was a result of a hot hand, it may just be because the probability of success was fluctuating throughout the course of the data generation or because we got the value of p wrong in our estimate. We also saw that variation in the length of the longest run was to be expected in data generated randomly, so we were not sure of what "too long" or "too short" should be in this process. In this section we will talk about random variables and their distributions so that we can bring greater clarity to such decision making processes. We will also talk about a commonly used test for randomness in two-valued sequences which does not require an estimate for the probability of success before use, see [2].

1. Random Variables

Definition 1.1. A Random Variable *is a rule that assigns a number to each outcome of an experiment. There may be more than one random variable associated with an experiment.*

Example 1.1. *An experiment consists of rolling a pair of dice, one red and one green, and observing the pair of numbers on the uppermost faces (red first). We let X denote the sums of the numbers on the uppermost faces. Below, we show the outcomes on the left and the values of X associated to some of the outcomes on the right:*

$$\begin{array}{cccccc}
\{(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)\}
\end{array}$$

Outcome	**X**
(1, 1)	2
(2, 1)	3
(3, 1)	4
(4, 1)	5
⋮	⋮

(a) What are the possible values of X?

Number of Runs In previous sections we thought about the length of runs of success and the length of runs of failures in Bernoulli trials separately. We can also consider the total number of runs in the data (including runs of both types) as a test statistic for non-randomness.

Example 1.2. *If we flip a coin 20 times giving us the following sequence of heads and tails:*

$$HHTTTHTTTTHHHTHHHHHH,$$

The number of runs is 7. Below, we show runs of tails in red and runs of Heads in black:

$$HH\color{red}{TTT}\color{black}{H}\color{red}{TTTT}\color{black}{HHH}\color{red}{T}\color{black}{HHHHHH},$$

Example 1.3. *An experiment consists of flipping a coin 4 times and observing the result ion sequence of heads and tails. The outcomes in the sample space are shown below.*

$$S.S. = \{HHHH,\ \ HHHT,\ \ HHTH,\ \ HHTT,\ \ HTHH,\ \ HTHT,\ \ HTTH,\ \ HTTT,$$
$$THHH,\ \ THHT,\ \ THTH,\ \ THTT,\ \ TTHH,\ \ TTHT,\ \ TTTH,\ \ TTTT\}.$$

(a) Let Y denote the number of runs observed. What are the possible values of Y?

(b) We could also define other variables associated to this experiment. Let X denote the number of heads observed. What are the possible values of X?

For some random variables, the possible values of the variable can be separated and listed in either a finite list or an infinite list. These variables are called **discrete random variables**.Some examples are shown below:

Experiment	**Random Variable, X**	**Possible values of X**
Roll a pair of six-sided dice	Sum of the numbers	$\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
Toss a coin 5 times	Number of tails	$\{0, 1, 2, 3, 4, 5\}$
Flip a coin until you get a tail	The number of coin flips	$\{1, 2, 3, \ldots, \}$
Flip a coin 50 times	Longest run of heads	$\{0, 1, 2, 3, \ldots, 50\}$

On the other hand, **a continuous random variable** can assume any value in some interval. Some examples are:

Experiment	**Random Variable, X**
Choose an NFL Quarterback at random	Height
Choose an NCAA Shot Putter at random	Arm Length
Choose a Track and Field athlete at random	Their best time for 100 meters

2. **Probability Distributions For Random Variables**

For a discrete random variable with finitely many possible values, we can calculate the probability that a particular value of the random variable will be observed by adding the probabilities of the outcomes of our experiment associated to that value of the random variable (assuming that we know those probabilities). This assignment of probabilities to each possible value of X is called the probability distribution of X.

Example 2.1. *If I roll a pair of fair six sided dice and observe the pair of numbers on the uppermost face, all outcomes are equally likely, each with a probability of* $\frac{1}{36}$*. Let* X *denote the sum of the pair of numbers observed. We saw that a value of 3 for* X *is associated to two outcomes in our sample space:* $(2, 1)$ *and* $(1, 2)$*. Therefore the probability that* X *takes the value 3 or* $P(X = 3)$ *is the sum of the probabilities of the two outcomes* $(2, 1)$ *and* $(1, 2)$ *which is* $\frac{2}{36}$*. That is*

$$P(X = 3) = \frac{2}{36}.$$

If X is a discrete random variable with finitely many possible values, we can display the probability distribution of X in a table where the possible values of X are listed alongside their probabilities.

Example 2.2. *I roll a pair of fair six sided dice and observe the pair of numbers on the uppermost face. Let* X *denote the sum of the pair of numbers observed. Complete the table showing the probability distribution of* X *below:*

$$\begin{array}{cccccc} \{(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)\} \end{array}$$

X	**P(X)**
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

Probability Distribution: If a discrete random variables has possible values $x_1, x_2, x_3, \ldots, x_k$, then a **probability distribution** $P(X)$ is a rule that assigns a probability $P(x_i)$ to each value x_i. More specifically,

- $0 \leq P(x_i) \leq 1$ for each x_i.
- $P(x_1) + P(x_2) + \cdots + P(x_k) = 1$.

Example 2.3. *An experiment consists of flipping a coin 4 times and observing the sequence of heads and tails. The random variable X is the number of heads in the observed sequence. The random variable Y is the length of the longest run of heads in the sequence and the random variable Z is the total number of runs in the sequence (of both H's and T's). Use the equally likely sample space*

$$S.S. = \{HHHH,\ HHHT,\ HHTH,\ HHTT,\ HTHH,\ HTHT,\ HTTH,\ HTTT,$$

$$THHH,\ THHT,\ THTH,\ THTT,\ TTHH,\ TTHT,\ TTTH,\ TTTT\}.$$

to fill in probabilities for each possible values of X, Y and Z in the tables below.

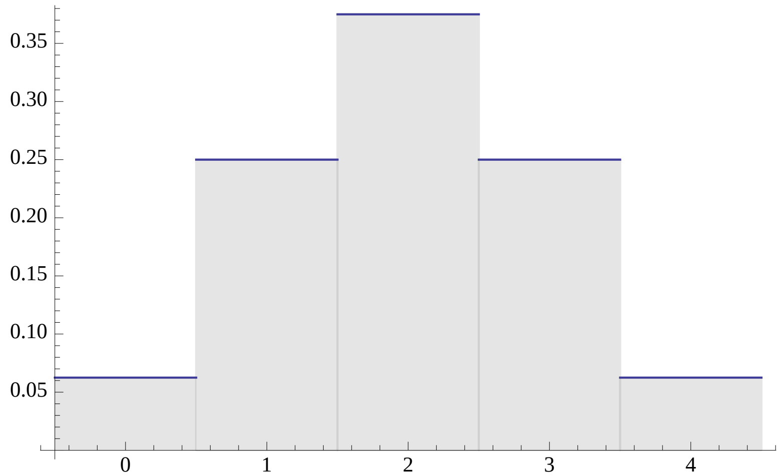
X (# Heads)	**P(X)**
0	
1	
2	
3	
4	

Y (longest Run of Heads)	**P(Y)**
1	
2	
3	
4	

Z (# Runs)	**P(Z)**
1	
2	
3	
4	

We can also represent a probability distribution for a discrete random variable with finitely many possible values **graphically** by constructing a bar graph. We form a category for each value of the random variable centered at the value which does not contain any other possible value of the random variable. We make each category of equal width and above each category we draw a bar with height equal to the probability of the corresponding value. if the possible values of the random variable are integers, we can give each bar a base of width 1.

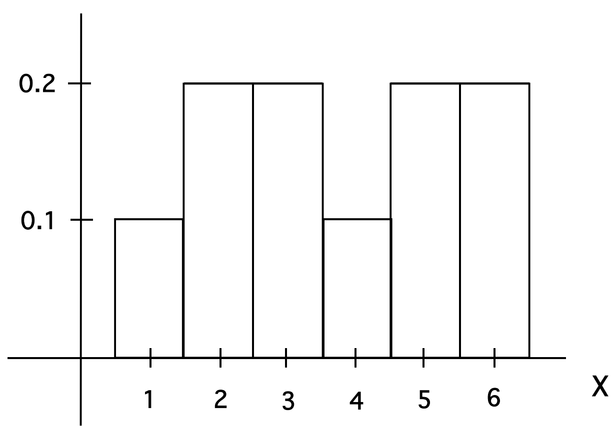
Example 2.4. *An experiment consists of flipping a coin 4 times and observing the sequence of heads and tails. The random variable X is the number of heads in the observed sequence. The following is a graphical representation of the probability distribution of X.*



Note We have already encountered a probability distribution and its plot in the previous section, when we talked about the probability distribution of the outcomes for experiment A_p and the random variable which measured the length of runs of baskets.

By Making all bars of equal width, we ensure that the graph adheres to the **area principle** in that the probability that any set of values will occur is equal to the area of the bars above those values. The total area of the distribution is 1.

Example The following is a probability distribution histogram for a random variable X.



What is $P(X \leq 5)$?

2.1. **Continuous Random Variables.** The probability distribution of a continuous random variable cannot be represented in a table since the possible values of the variable cannot be separated. The distribution is represented using the graphical method as a continuous curve and is called a probability density function. Probabilities are calculated for intervals instead of particular values. The probability that the value of a random variable will fall in the interval $[a, b]$, denoted $P(a \leq X \leq b)$ is given by the area under the probability density function above that interval. The area under the entire probability density curve is 1.

Example 2.5.

The picture below taken form the website height-differences-among-professional-athletes. It shows three probability density functions for the height (in inches) of NFL, NBA and NHL players respectively(Compiled in November 2009).

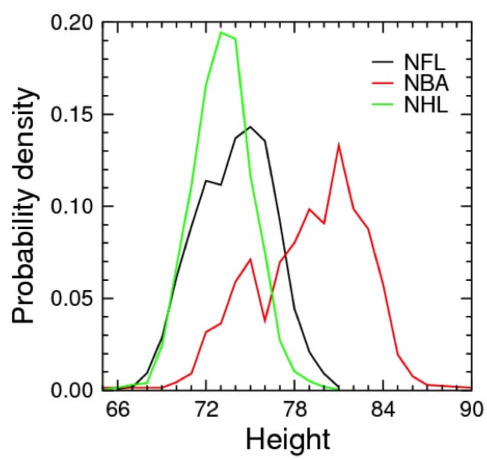


Fig. 1: Height distribution of all players in the NFL, NBA, and NHL.

(a) If we were to choose an NFL player at random from the group and observe his height (in 2009), the probability that his height will be above 77 inches corresponds to the area under the curve corresponding to the NFL players to the right of 77 inches. Shaded the region under the curve above corresponding to this probability and make a rough guess of the probability.

3. **Expected Value of a Random Variable**

Consider the above experiment where we flip a coin 4 times and graph the probability distribution for the random variable X which is equal to the number of heads in the sequence. The graph of the distribution is symmetric and centered at the value $X = 2$. This is the balance point of the graph in that if we were to balance the graph on a single point of support, this is where we would place the support. This balance point gives us a measure of the center of the distribution. It is a very important and widely used statistic and is related to our expectations for the average value of the variable over many trials of the experiment.

3.1. **Average of a set of observations.** The following sequence is the result of 10 trials of the experiment "roll a fair six sided die 10 times".

$$1, 6, 3, 2, 5, 2, 3, 2, 4, 6, 4, 6, 2, 6, 3, 6, 2, 6, 3, 5.$$

The random variable recorded is X equal to the number appearing on the uppermost face. We can calculate the average of this data by adding all of the numbers and dividing by 20 (we denote the mean from this sample of data by $\bar{x}$). We can make our work easier by making a frequency table showing each outcome alongside the frequency with which it occurs in the data:

Outcome i	**Frequency** f_i
1	1
2	5
3	4
4	2
5	2
6	6

$\bar{x} =$ ____

We can write a formula for the mean as

$$\frac{1f_1 + 2f_2 + 3f_3 + 4f_4 + 5f_5 + 6f_6}{20},$$

where the frequency of the outcome "i" in the data is denoted by f_i. We can also perform this calculation as the sum of the outcomes times their relative frequencies

$$1\frac{f_1}{20} + 2\frac{f_2}{20} + 3\frac{f_3}{20} + 4\frac{f_4}{20} + 5\frac{f_5}{20} + 6\frac{f_6}{20}.$$

In a small set of data, we do not expect the relative frequencies to match the probabilities, but in a very large set of data, we would expect each of the relative frequencies to match the probability of the outcome which is 1/6 for all of the outcomes in this case. Thus for a large number (N) of trials of this experiment, we would expect the average of the numbers recorded to be approximately

$$1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = 3.5.$$

This number is called the expected value of X and is denoted by $E(X)$ or μ.

3.2. **Expected Value of a Discrete Random Variable.**

Definition 3.1. *If X is a random variable with a finite number of possible values $x_1, x_2, \ldots, x_n$ and corresponding probabilities $p_1, p_2, \ldots, p_n$, the* **expected value of** *X, denoted by $E(X)$ or μ, is*

$$\boxed{\mu = E(X) = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n.}$$

Outcomes X	**Probability P(X)**	**Out. × Prob. XP(X)**
x_1	p_1	$x_1 p_1$
x_2	p_2	$x_2 p_2$
$\vdots$	$\vdots$	$\vdots$
x_n	p_n	$x_n p_n$
		Sum $= E(X) = \mu$

If we run a **large number of trials of the experiment**, say N, and observe the value of the random variable X in each, $x_1, x_2, x_3, \ldots, x_N$, we should have that $E(X) \approx \frac{x_1+x_2+x_2+\cdots+x_N}{N}$ or

$$\boxed{E(X)N \approx x_1 + x_2 + x_2 + \cdots + x_N.}$$

Graphically this number corresponds to the **balance point of the distribution** and we expect the observations of the variable X to average to $E(X)$ over many trials of the experiment. if the variable is discrete but has infinitely many possible outcomes, we can use infinite summation to calculate the expected value, however this is beyond the scope of this course.

Example 3.1. *An experiment consists of flipping a coin 4 times and observing the sequence of heads and tails. The random variable X is the number of heads in the observed sequence. The random variable Y is the length of the longest run of heads in the sequence and the random variable Z is the total number of runs in the sequence (of both H's and T's). Find the expected value of each random variable using their probability distributions shown below*

X	**P(X)**
0	
1	
2	
3	
4	

Y	**P(Y)**
1	
2	
3	
4	

Z	**P(Z)**
1	
2	
3	
4	

3.3. **Expected Value of a Continuous Random Variable.** For a continuous random variable, X, we can use a method from calculus called integration to calculate the expected value. This is beyond the scope of this course, but $E(X)$ can be thought of geometrically as the balance point of the probability density function in this case. As above, we can also interpret the expected value, $E(X)$, as the number we would expect to get if we calculated the average of the observations of the variable X over many trials of the experiment.

Example 3.2. *Estimate the average height of an NHL player from the picture of the probability density function shown below by estimating the balance point:*

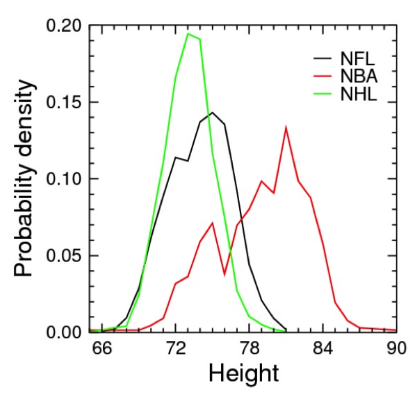


Fig. 1: Height distribution of all players in the NFL, NBA, and NHL.

4. Standard Deviation of a Random Variable

If we run an experiment such as choosing an NHL player at random from the population of NHL players shown above, and observe their height, we would not be surprised if the height of the player chosen was not exactly the expected value (approximately 73 inches). We expect some variability in the outcomes of this experiment. However from looking at the diagram, we see that it is unlikely that the player chosen would have height more than 3.5 inches more or less than the expected value. On the other hand, if we chose an NBA player at random from the population shown above, we would not be surprised if the player was 5 inches above or below the average (approx. 79 in.). The distribution of heights in the NBA is more spread out than that in the NHL leading us to expect more variability in the data.

A commonly used **measure of variability** in distributions of random variables is the **standard deviation**. The square of the standard deviation, called the **variance**, is also commonly used. The variance of a random variable can be viewed as the average squared distance from the mean or expected value of a random variable over a large number of trials of the experiment. The standard deviation is the square root of the variance. This may be viewed as an attempt to create a measure of deviation from the mean in the original units in which the variable is measured.

Variance and standard deviation of a random variable If we roll a fair six sided die and observe the number on the uppermost face (X), the (expected) average squared distance of the outcomes from $E(X) = 3.5$ in a large set of data is given by the sum of the squared distance from 3.5 for each outcome times its probability.

Outcome	Probability
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

To find the variance (the (expected)average squared distance from the mean, $\mu = 3.5$) one would estimate that over many trials of this experiment the **variance** denoted $Var(X)$, or σ^2, would be roughly

$$Var(X) = \frac{1}{6}\times(1-3.5)^2+\frac{1}{6}\times(2-3.5)^2+\frac{1}{6}\times(3-3.5)^2+\frac{1}{6}\times(4-3.5)^2+\frac{1}{6}\times(5-3.5)^2+\frac{1}{6}\times(6-3.5)^2 = \frac{35}{12}$$

The **standard deviation** of X, denoted by σ or $\sigma(X)$, is the square root of the variance, $\sigma = \sqrt{\sigma^2}$. In this case $\sigma = \sqrt{\frac{35}{12}} \approx 1.71$.

Definition 4.1. *If X is a random variable with values $x_1, x_2, \ldots, x_n$, corresponding probabilities $p_1, p_2, \ldots, p_n$, and expected value $\mu = E(X)$, then*

$$\boxed{\textbf{Variance} = \sigma^2(X) = p_1(x_1-\mu)^2 + p_2(x_2-\mu)^2 + \cdots + p_n(x_n-\mu)^2}$$

and

$$\boxed{\textbf{Standard Deviation} = \sigma(X) = \sqrt{Variance}}.$$

$\mathbf{x_i}$	$\mathbf{p_i}$	$\mathbf{x_i p_i}$	$(\mathbf{x_i} - \mu)$	$(\mathbf{x_i} - \mu)^\mathbf{2}$	$\mathbf{p_i}(\mathbf{x_i} - \mu)^\mathbf{2}$
x_1	p_1	x_1p_1	$(x_1-\mu)$	$(x_1-\mu)^2$	$p_1(x_1-\mu)^2$
x_2	p_2	x_2p_2	$(x_2-\mu)$	$(x_2-\mu)^2$	$p_2(x_2-\mu)^2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	p_n	x_np_n	$(x_n-\mu)$	$(x_n-\mu)^2$	$p_n(x_n-\mu)^2$
		Sum $= \mu$			**Sum** $= \sigma^2(X)$

Example 4.1. *An experiment consists of flipping a coin 4 times and observing the sequence of heads and tails. The random variable Z is the number of runs in the sequence. Find $E(Z)$ and the standard deviation, $\sigma(Z)$*

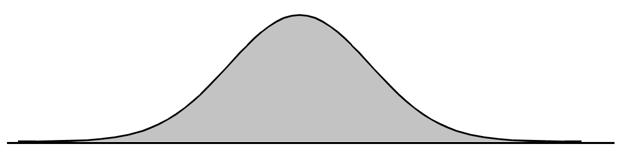
Z	**P(Z)**
1	7/16
2	6/16
3	6/16
4	2/14

5. **Interpreting The Standard Deviation**

When presented with raw scores for performance, it is difficult to interpret their meaning without some measure of center and variability for the population from which they come. In any set of data, whether it is population data or a sample, observations that are more than 3 standard deviations from the mean are rare and exceptional. One such rule demonstrating this is the empirical rule for bell shaped data shown below.

5.1. **The Empirical Rule for bell Shaped Distributions.**

The Bell curve or normal distribution frequently occurs as a population distribution for continuous variables. Also sometimes the probability distribution for a discrete random variable can be well approximated by a bell shaped curve and one can use the bell curve to approximate probabilities for the discrete variable. The empirical rule given below applies to random variables with probability distributions that are bell shaped, like the one shown below.



Empirical Rule If a random variable has a probability distribution which is bell shaped or approximately bell shaped, we have the following empirical rule:

- The probability of getting an outcome within one standard deviation of the mean on any given trial of the experiment is approximately 0.68. That is $P(\mu - \sigma, \mu + \sigma) \approx 0.68$.
- The probability of getting an outcome within two standard deviations of the mean on any given trial of the experiment is approximately 0.95. That is $P(\mu - 2\sigma, \mu + 2\sigma) \approx 0.95$.
- The probability of getting an outcome within three standard deviations of the mean on any given trial of the experiment is approximately 0.997. That is $P(\mu - 3\sigma, \mu + 3\sigma) \approx 0.997$.

Bell shaped distributions are very important because they frequently occur as population distributions. Even more importantly, the central limit theorem says that if we take all samples of a given size from a population and calculate all of the means, then the distribution of the means is bell shaped (Normal).

Numerical Measures of Relative Standing

Quite often when interpreting a data observation, such as a baby's height and weight, we are interested in how it compares to the rest of the relevant population. Measures of relative standing describe the location of a particular measurement relative to the rest of the data. We explore some of the standard measures of relative standing below.

Z-Scores The z-score for a particular measurement in a set of data, measures how many standard deviations that measurement lies away from the mean.

The **z-score** for a data measurement, x, of the random variable X, resulting from a trial of an experiment is

$$z = \frac{x - \mu}{\sigma}$$

where $\mu = E(X)$ and $\sigma = \sigma(X)$.

Example 5.1. *The NFL combine is a week-long showcase where college football players perform physical and mental tests in front of National Football League coaches, general managers, and scouts. The following table shows the mean and standard deviation for that population for each physical test for the players in each category. On this* `webpage` *the statistics for the performance of the top 750 prospects for the NFL draft at the NFL yearly combine over a 7 year period from 2005 to 2011. From the graph shown for the 40-yard dash, we see that the mean time for the wide receivers among these players is approximately 4.51 seconds and the standard deviation is roughly 0.1 seconds. From the table shown, we see that the mean time for the cone test for wide receivers is 6.96 with a standard deviation of 0.2.*

One of our wide receivers at Notre Dame, Golden Tate, participated in the 2010 NFL combine. His time for the 40-yard dash was 4.42 seconds and for the cone drill, his time was 7.12 seconds. In which skill test did he have the better z-score (as a member of the above population)?

6. Wald Wolfowitz Runs Test

Now that we have some more sophisticated tools at our disposal for making decisions, let's review our exploration of the question of the hot hand. One random variable we used for comparison to decide if a sequence of success' and failures was generated randomly was the length of the longest run in the data. By now you probably realize that we could make better decisions using this variable as our test variable if we knew something about its variance and standard deviation. This topic is beyond the scope of this course but a nice discussion of it can be found in Schilling [1]. We will however have a look at a different statistic, namely the number of runs in a set of data with two values discussed in Wald and Wolfowitz [2].

Given a sequence with two values, success (S) and failure (F), with N_s success' and N_f failures, let X denote the number of runs (of both S's and F's). Wald and Wolfowitz determined that for a random sequence of length N with N_s success' and N_f failures (note that $N = N_s + N_f$), the number of runs has mean and standard deviation given by

$$E(X) = \mu = \frac{2N_sN_f}{N} + 1, \qquad \sigma(X) = \sqrt{\frac{(\mu-1)(\mu-2)}{N-1}}.$$

The distribution of X is approximately normal if N_s and N_f are both bigger than 10.
(see `http://www.itl.nist.gov/div898/handbook/eda/section3/eda35d.htm` for more details).

Now we can use this information to test if a sequence is likely to have been generated randomly (as a sequence of independent identical Bernoulli trial) or not. If the number of runs in the sequence is too far (more than 3 standard deviations) from what is expected, we could conclude that it is highly unlikely that the sequence was generated randomly. (This is a rough description of the concept of Hypothesis Testing which you can learn about in most statistics courses). To run the test, we calculate the observed value of X (the number of runs for the given sequence) denoted by x below. We calculate the values of N, N_s and N_f from our sequence and if the Z- value:

$$Z = \frac{x-\mu}{\sigma} \;=\; \frac{x - \left(\frac{2N_sN_f}{N}+1\right)}{\sqrt{\frac{(\mu-1)(\mu-2)}{N-1}}}.$$

If $Z > 3$ or $Z < -3$ we decide that the sequence was not generated randomly. (The chances we are wrong in our decision is roughly 0.003 by the empirical rule.) If we decide that the sequence was not generated randomly, it does not necessarily imply that there is a hot hand effect, it may just be that the probability of success varies from trial to trial in a pattern different from that of the hot hand. If on the other hand we observe a value of Z which has absolute value less than or equal to 3, it does not mean that the sequence was generated randomly, rather it means that there is insufficient evidence for us to reject the idea.

Example 6.1. *We would expect that the following sequences of Heads (H) and Tails (T) are unlikely to fit the profile of sequences generated randomly in a series of independent identical Bernoulli Trials:*
Seq 1: $H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T$
Seq 2: $H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ T\ T\ T\ T\ T\ T\ T\ T\ T\ T\ T\ T\ T\ T\ T$
Seq 3: $H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ T\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H\ H$

Use the Walf Wolfowitz Runs Test to test for randomness.

Example 6.2. *Use the Wald Wolfowitz Runs Test to test the following sequence of consecutive baskets and misses for basketball player J R Smith for randomness:*

MMBBMMMMBMMMBBMMBMBMMBMMBMBMMBBMMMMBBMBMBBBBMMBMBMMBBBBMMM

References

1. Mark F. Schilling, *The surprising predictability of long runs*, Math. Mag. **85** (2012), no. 2, 141–149. MR 2910308
2. A. Wald and J. Wolfowitz, *On a test whether two samples are from the same population.*, Ann. Math. Statist. **11** (1940), 147–162.