

# CALCULUS AND PROBABILITY BASED STATISTICS

MARK ANDREW GERADS

## 1. FINITE STATISTICS

**1.1. Introduction.** The purpose of this text is to have statistics built on a solid foundation of probability and calculus.

**Definition 1.1.**  $A$  is an independently repeatable event.

**Definition 1.2.**  $p \in [0, 1]$  is the probability that  $A$  will happen when  $A$  is given a chance to happen. For simplicity, we assume that  $p$  is the same every time  $A$  is given a chance to happen.

**Definition 1.3.**  $n \in \mathbb{W}$  is the number of times  $A$  had a chance to happen.

**Definition 1.4.**  $k \in \mathbb{W}$  is the number of times  $A$  actually happened.

*Remark 1.5.*  $k \leq n$ .

**Definition 1.6.**  $f_p$  is the probability distribution for  $p$ .

*Remark 1.7.*  $1 = \int_0^1 f_p dp$

*Remark 1.8.*  $k/n$  is the expected probability of  $p$  and is the maximum value of the function  $f_p$ .

**Definition 1.9.**  $(a, b)$  is the confidence interval of  $p$  such that  $b - a$  is minimized.

**Definition 1.10.**  $c$  is the desired certainty for the confidence interval of  $p$ .

*Remark 1.11.*  $0 \leq c = \int_a^b f_p dp \leq 1$

*Remark 1.12.*  $0 \leq a \leq b \leq 1$

**1.2. The Problem.**  $n, k, c$  are known, and we must solve for  $a, b$ .

*Remark 1.13.* The probability of  $k$  events given  $n$  chances is  $\binom{n}{k} p^k (1-p)^{n-k}$ , but we must scale it to satisfy 1.7:

$$(1.1) \quad f_p = \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\int_0^1 \binom{n}{k} p^k (1-p)^{n-k} dp} = \frac{p^k (1-p)^{n-k}}{\int_0^1 p^k (1-p)^{n-k} dp} = (n+1) \binom{n}{k} p^k (1-p)^{n-k}$$

**1.3. The 4 cases of finding the interval.**

**1.3.1. Case  $n = k = 0$ .** This case is self-symmetrical.

*Remark 1.14.* We have

$$(1.2) \quad (a, b) = \left( \frac{1-c}{2}, \frac{1+c}{2} \right)$$

1.3.2. *Case*  $n > k = 0$ . This case is symmetrical to case 1.3.3.

*Remark 1.15.* We can solve

$$(1.3) \quad \int_0^b (n+1)(1-p)^n dp = c = 1 - (1-b)^{n+1}$$

*Remark 1.16.* We have

$$(1.4) \quad (a, b) = \left(0, 1 - (1-c)^{\frac{1}{n+1}}\right)$$

**Example 1.17.** Now we can analyze the old saying "Third time's the charm.".  $n = 2$ .

$$(1.5) \quad c = 0.5 \Rightarrow b \approx 0.2062994740, c = 0.95 \Rightarrow b \approx 0.6315968501$$

We can now see that if you fail at something twice, then success before the 8th try is worth approximately a coin flip.

**Example 1.18.** Suppose we were repeatedly trying a difficult task. How many times of trying without success must there be for us to conclude with  $c = 0.99$  that  $b \leq 0.01$ ?

$$(1.6) \quad n = \frac{2 \log(3) + \log(11)}{2 \log(2) - 2 \log(3) + 2 \log(5) - \log(11)} \approx 457.2105766$$

We can see "If at first you don't succeed, try, try again." is sensible for things that we want to be certain are improbable before giving up.

1.3.3. *Case*  $n = k > 0$ . This case is symmetrical to case 1.3.2.

*Remark 1.19.* We can solve

$$(1.7) \quad \int_a^1 (k+1)p^k dp = c = 1 - a^{k+1}$$

*Remark 1.20.* We have

$$(1.8) \quad (a, b) = \left((1-c)^{\frac{1}{k+1}}, 1\right)$$

1.3.4. *Case*  $n > k > 0$ . This case is self-symmetrical.

*Remark 1.21.* We desire to solve

$$(1.9) \quad \int_a^b (n+1) \binom{n}{k} p^k (1-p)^{n-k} dp = c = (n+1) \binom{n}{k} (B_b(k+1, -k+n+1) - B_a(k+1, -k+n+1))$$

where  $B$  is the incomplete beta function.

*Remark 1.22.* We can use the hints

$$(1.10) \quad a^k (1-a)^{n-k} = b^k (1-b)^{n-k}$$

*Remark 1.23.* By solving

$$(1.11) \quad \frac{\partial f_p}{\partial p} = 0$$

, we get

$$(1.12) \quad \lim_{c \rightarrow 0} (a, b) = \left(\frac{k}{n}, \frac{k}{n}\right), m = (n+1) \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k}$$

1.4. Table of small trial half probability intervals. Table with  $c = \frac{1}{2}$ :

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$n = 0$	$a \approx 0.2500000000$ $b \approx 0.7500000000$				
$n = 1$	$a \approx 0.0000000000$ $b \approx 0.2928932188$	$a \approx 0.7071067812$ $b \approx 1.0000000000$			
$n = 2$	$a \approx 0.0000000000$ $b \approx 0.2062994740$	$a \approx 0.326351822$ $b \approx 0.673648178$	$a \approx 0.7937005260$ $b \approx 1.0000000000$		
$n = 3$	$a \approx 0.0000000000$ $b \approx 0.1591035847$	$a \approx 0.1966713241$ $b \approx 0.4921993002$	$a \approx 0.507800700$ $b \approx 0.803328676$	$a \approx 0.8408964153$ $b \approx 1.0000000000$	
$n = 4$	$a \approx 0.0000000000$ $b \approx 0.1294494367$	$a \approx 0.1387754454$ $b \approx 0.3894044243$	$a \approx 0.3594361648$ $b \approx 0.640563835$	$a \approx 0.610595575$ $b \approx 0.861224555$	$a \approx 0.8705505633$ $b \approx 1.0000000000$
$n = 5$	$a \approx 0.0000000000$ $b \approx 0.1091012819$	$a \approx 0.1065967785$ $b \approx 0.3226910840$	$a \approx 0.2770309506$ $b \approx 0.532125995$	$a \approx 0.4678740055$ $b \approx 0.722969050$	$a \approx 0.677308916$ $b \approx 0.893403222$
$n = 6$	$a \approx 0.0000000000$ $b \approx 0.09427633574$	$a \approx 0.0862905298$ $b \approx 0.2757178692$	$a \approx 0.2249291772$ $b \approx 0.4551185682$	$a \approx 0.3788484407$ $b \approx 0.621151559$	$a \approx 0.544881432$ $b \approx 0.775070823$
$n = 7$	$a \approx 0.0000000000$ $b \approx 0.08299595680$	$a \approx 0.07237321411$ $b \approx 0.2407853463$	$a \approx 0.1891248358$ $b \approx 0.3976355162$	$a \approx 0.3180913474$ $b \approx 0.544067630$	$a \approx 0.4559323705$ $b \approx 0.681908652$
$n = 8$	$a \approx 0.0000000000$ $b \approx 0.07412528771$	$a \approx 0.06226629927$ $b \approx 0.2137614499$	$a \approx 0.1630534052$ $b \approx 0.3530822174$	$a \approx 0.2740192600$ $b \approx 0.4838784786$	$a \approx 0.3919638426$ $b \approx 0.608036157$
	$k = 5$	$k = 6$	$k = 7$	$k = 8$	
$n = 5$	$a \approx 0.8908987181$ $b \approx 1.0000000000$				
$n = 6$	$a \approx 0.724282131$ $b \approx 0.913709470$	$a \approx 0.9057236643$ $b \approx 1.0000000000$			
$n = 7$	$a \approx 0.602364484$ $b \approx 0.810875164$	$a \approx 0.759214654$ $b \approx 0.927626786$	$a \approx 0.9170040432$ $b \approx 1.0000000000$		
$n = 8$	$a \approx 0.516121520$ $b \approx 0.725980740$	$a \approx 0.646917782$ $b \approx 0.836946595$	$a \approx 0.786238550$ $b \approx 0.937733701$	$a \approx 0.9258747123$ $b \approx 1.0000000000$	