

Warm up 11:00 - 11:10

Stats 134

Chapter 2

1. Ten cards are dealt off the top of a well shuffled deck. The binomial formula doesn't apply to find the chance of getting exactly three diamonds because:
 - a The probability of a trial being successful changes
 - b** The trials aren't independent
 - c There isn't a fixed number of trials
 - d more than one of the above

The (unconditional) probability of getting a diamond is always $\frac{1}{4}$. However the trials are dependent on one another,

$$X = \text{\# diamonds in 10 trials}$$

$$P(X=3) = \binom{10}{3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^7$$

Last time

Sec 2.1 The Binomial Distribution

A Bernoulli trial has 2 outcomes, success and failure.

Prob P

Prob $q = 1 - P$

(think of tossing a coin having prob P of landing head)

n independent Bernoulli trials, each with prob P of success, has a Binomial distribution written $\text{Bin}(n, p)$.

For $k = 0, 1, 2, \dots, n$ the Binomial formula

$$\text{say } P(k) = \binom{n}{k} P^k q^{n-k}$$

↑
the prob
of getting exactly
 k out of n successes.
n!
K!(n-K)!

if $n=5, P=\frac{1}{2}$

$$P(0) = \binom{5}{0} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

$$P(1) = \binom{5}{1} \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 = \frac{5}{32}$$

$$P(2) = \binom{5}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 = \frac{10}{32}$$

$$P(3) = \frac{10}{32}$$

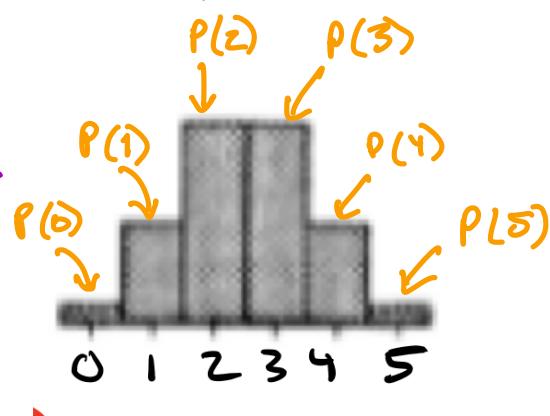
$$P(4) = \frac{5}{32}$$

$$P(5) = \frac{1}{32}$$

number of arrangements of two 1s in 5 slots

$$\binom{5}{2} = \frac{5!}{2! \cdot 3!} = 10$$

1	1	0	0	0
1	0	1	0	0
1	0	0	1	0
1	0	0	0	1
0	1	1	0	0
0	1	0	1	0
0	1	0	0	1
0	0	1	1	0
0	0	1	0	1
0	0	0	1	1



Ex

Halloween will be here before you know it and the children in your neighborhood will come trick-or-treating (that is, they will come to your door and demand candy). Suppose there are 20 children in your neighborhood and 30 houses (one of which is yours). Each child independently chooses 10 houses at random without replacement to visit.

- a What is the probability that a specific child will visit your house? $P = \frac{1}{3}$
b What is the probability that exactly 10 children visit your house?

$$n = 20 \quad P = \frac{1}{3}$$
$$k = 10$$
$$\left[\binom{20}{10} \left(\frac{1}{3}\right)^{10} \left(\frac{2}{3}\right)^{10} \right]$$

Todays

- ① Finish sec 2.1 Binomial distributions
- ② Start sec 2.2 Normal approximations to the binomial.

(1) Binomial Dist. The mode and mean,
 what does the binomial distribution look like
 for different n, p ?
 (most likely outcome)

Def'n (Mode)

The mode of the binomial distribution is the most likely outcomes (i.e. the k such that $P(k) = \binom{n}{k} p^k (1-p)^{n-k}$ is largest)

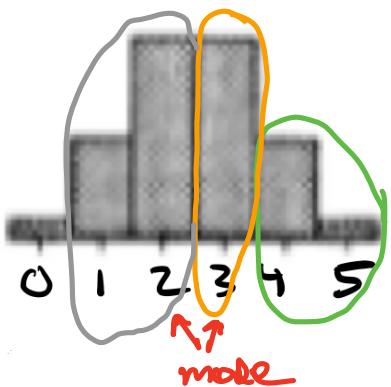
* Proof at end of notes using, $\frac{P(k)}{P(k-1)}$, the consecutive odds ratio,
 Then For $k \in \{1, 2, \dots, n\}$

- $K < np + p$ iff $P(k-1) < P(k)$
- $K > np + p$ iff $P(k-1) > P(k)$
- $K = np + p$ iff $P(k-1) = P(k)$

given by Binomial formula,

Picture

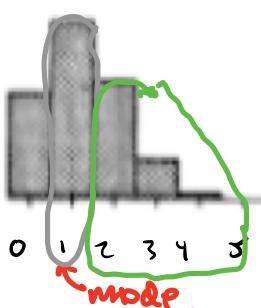
ex



$$\begin{aligned} n &= 5 \\ k &= 0, 1, 2, 3, 4, 5 \\ p &= 1/2 \end{aligned}$$

$$np + p = 5(1/2) + 1/2 = 3$$

ex



$$\text{ex } n=5, p=1/4, k=0, 1, 2, 3, 4, 5$$

$$np + p = 5 \cdot \frac{1}{4} + \frac{1}{4} = 1.5$$

The mode for the binomial distribution has 2 cases:

$$\text{mode} = \begin{cases} m & \text{if } np+p \notin \mathbb{Z} \\ m+1, m & \text{if } np+p \in \mathbb{Z} \end{cases}$$

integers.

where $m = \lfloor np+p \rfloor$

↑
the int ger
part o
f $np+p$

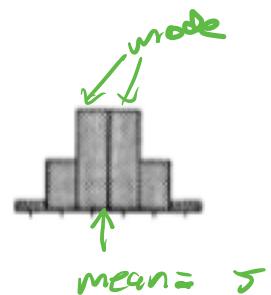
The mode is a measure of the center of the data.

However, the true center of your data is the expectation (i.e the mean).

Fact ← shown in Chap 3
The expected (mean) number of success is
 $\Rightarrow \boxed{\mu = np}$

This isn't usually an integer

$$\text{e.g. } n=5 \quad \mu = np = 5/2 \\ p=1/2$$



Given $np \neq 1$, if the mean $= \frac{np}{2}$ is an integer is it the mode?

Yes, $np \in \mathbb{Z} \Rightarrow np+p \notin \mathbb{Z} \Rightarrow$ there is a single mode $m = \lfloor np+p \rfloor = np$

↑
the mode

→ shown in Chap 3

Fact the average spread around the mean

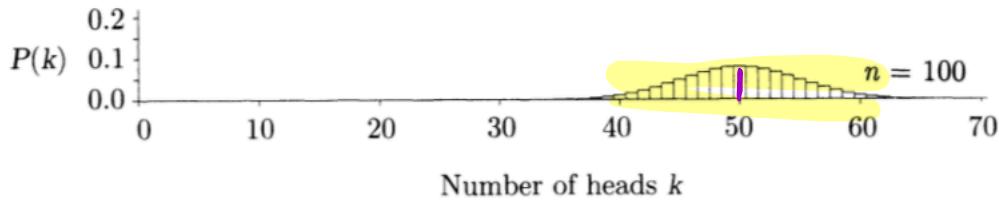
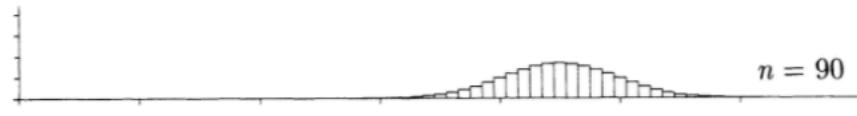
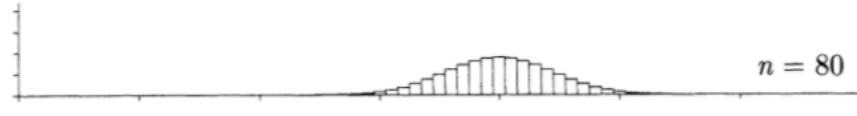
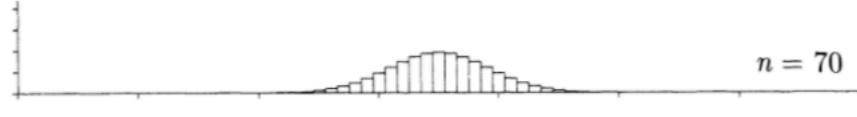
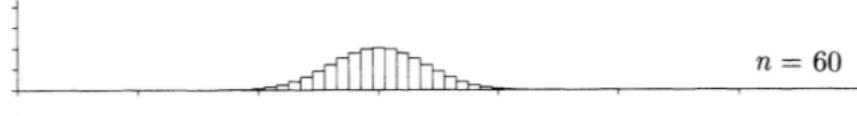
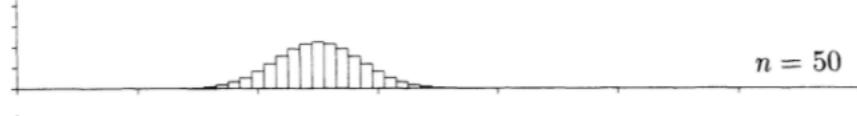
(standard deviation) is $\sigma = \sqrt{npq}$ where $q = 1-p$

Notice that the spread around the mean gets larger as n gets bigger.

P88

Binomial ($n, \frac{1}{2}$)

$$\sigma = \sqrt{n p q} = \sqrt{n \frac{1}{2} \cdot \frac{1}{2}} = \frac{\sqrt{n}}{2}$$



Number of heads k

law of averages say ↴

$X = \# \text{ heads}$

$\frac{X}{n} \rightarrow P = \frac{1}{2}$

Stat 134

1. A fair coin is tossed, and you win a dollar if there are exactly 50% heads. Which is better?

a 10 tosses

b 100 tosses

$$n=10 \quad P(5) = \binom{10}{5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^5 = .246$$

$$n=100 \quad P(50) = \binom{100}{50} \left(\frac{1}{2}\right)^{50} \left(\frac{1}{2}\right)^{50} = .08$$

② sec 2.2 The normal distribution

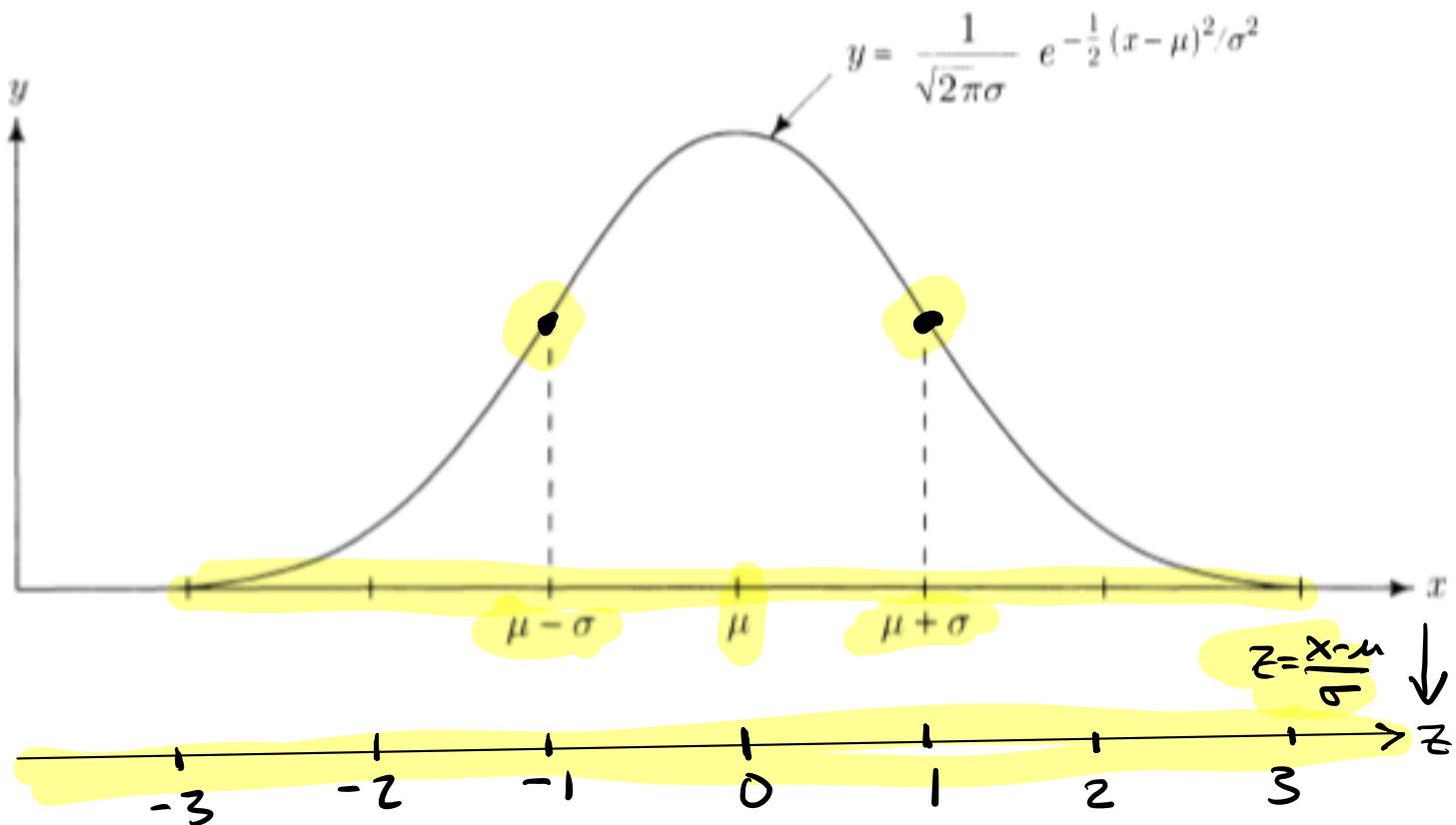
The normal curve is $y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Notice :

- ① two param $\mu = \text{mean}$
 $\sigma = \text{std dev}$
- ② inflection pts $\mu \pm \sigma$
- ③ almost all data between $\mu \pm 3\sigma$

P93

FIGURE 1. The normal curve.



To find the area under the curve it is convenient to make a change of coordinates

$$z = \underline{x - \mu}$$

This makes $\mu=0$ and $\sigma=1$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

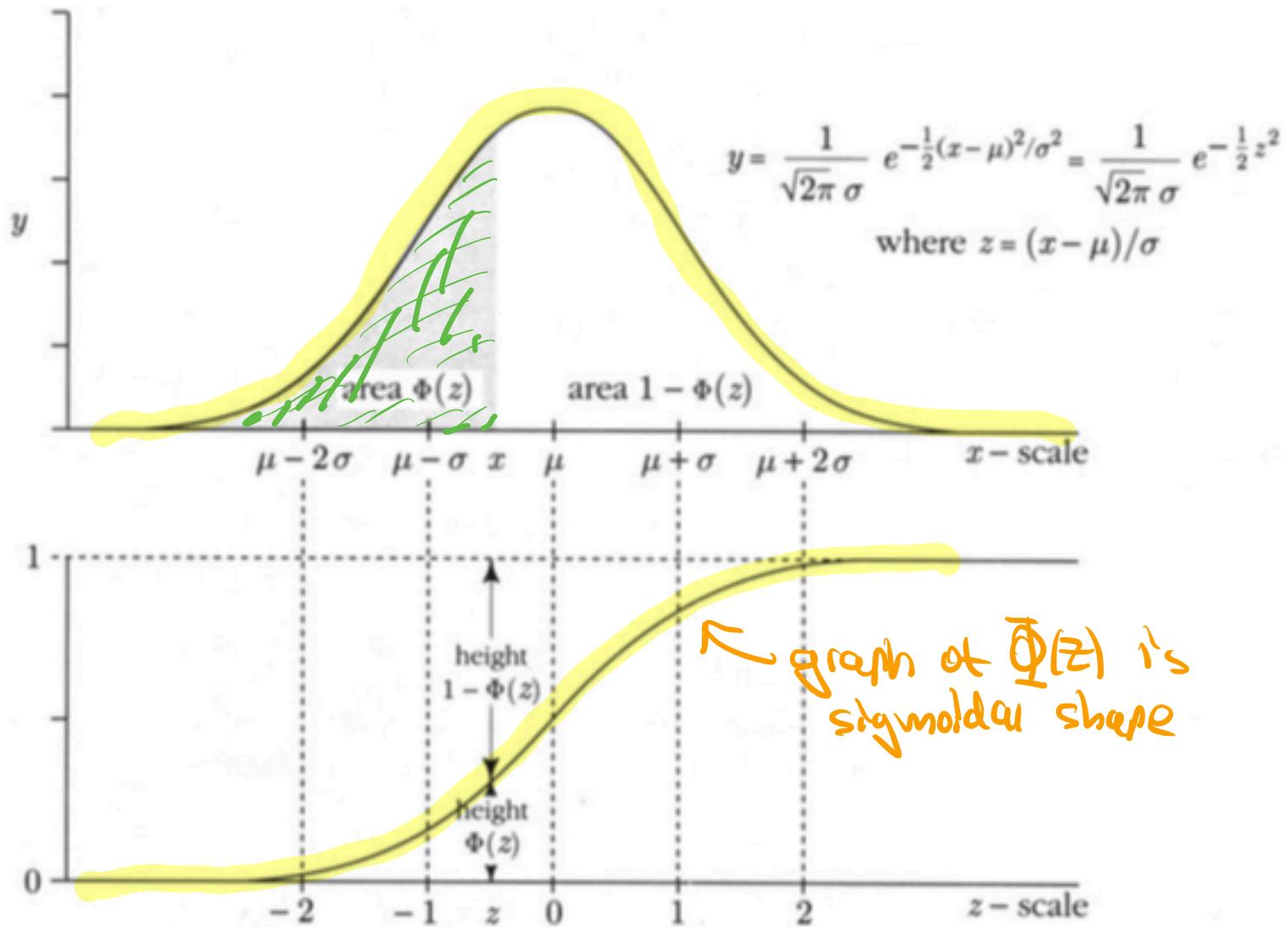
std normal curve

Define cumulative distribution function (cdf)

as $\Phi(z) = \int_{-\infty}^z \phi(t) dt$

\leftarrow area between $-\infty$ and z

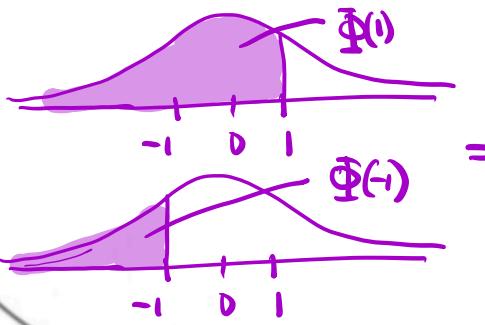
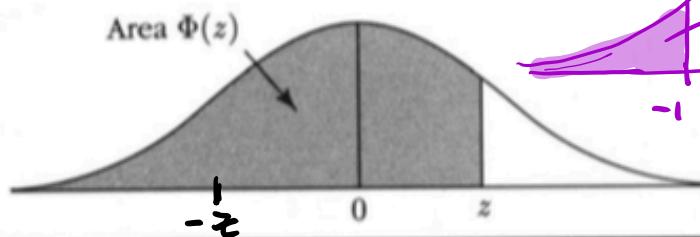
P95



We can't solve integral $\Phi(z)$ but instead use a look up table.

Notice

$$\Phi(-z) = 1 - \Phi(z)$$



Appendix 5

Normal Table

Table shows values of $\Phi(z)$ for z from 0 to 3.59 by steps of .01. Example: to find $\Phi(1.23)$, look in row 1.2 and column .03 to find $\Phi(1.2 + .03) = \Phi(1.23) = .8907$. Use $\Phi(z) = 1 - \Phi(-z)$ for negative z .

Find area between 1 and -1 in std normal curve:

$$\begin{aligned} & \Phi(1) - \bar{\Phi}(-1) \\ &= \bar{\Phi}(1) - (1 - \bar{\Phi}(1)) \\ &= 2\bar{\Phi}(1) - 1 \end{aligned}$$

$$= 2(.8413) - 1$$

$$= .68$$

Find area between z and $-z$.

$$2\bar{\Phi}(z) - 1$$

$$= 2(.9772) - 1$$

$$= .95$$

Find area between 3 and -3.

$$.997$$

	.0	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

This is known as the empirical rule,

Appendix

Thm

- (1) $K < np + p$ iff $P(K-1) < P(K)$
- (2) $K > np + p$ iff $P(K-1) > P(K)$
- (3) $K = np + p$ iff $P(K-1) = P(K)$

PF/

First note that $\frac{\binom{n}{K}}{\binom{n}{K-1}} = \frac{\frac{n!}{K!(n-K)!}}{\frac{n!}{(K-1)!(n-K+1)!}} = \boxed{\frac{n-K+1}{K}}$

Called the consecutive odds ratio

$\frac{P(K)}{P(K-1)}$ / binomial formula $= \frac{\binom{n}{K} p^K (1-p)^{n-K}}{\binom{n}{K-1} p^{K-1} (1-p)^{n-K+1}} = \boxed{\frac{n-K+1}{K} \cdot \frac{p}{1-p}}$

$$P(K-1) < P(K)$$

$$\Leftrightarrow 1 < \frac{P(K)}{P(K-1)}$$

$$\Leftrightarrow 1 < \frac{n-K+1}{K} \cdot \frac{p}{1-p}$$

$$\Leftrightarrow K(1-p) < (n-K+1)p$$

$$\Leftrightarrow K - kp < np - nk + p$$

$$\Leftrightarrow K < np + p$$

$$\text{so } P(K-1) < P(K) \Leftrightarrow K < np + p$$

similarly for $>$ or $=$ instead of $<$

□