Binomial distribution (consider symmetric case for simplicity)

n flips, mean $\mu = n/2$, s.d. $\sigma = \frac{1}{2}\sqrt{n}$

binomial approaches Normal

if $x = \#$ heads; $\quad X = \dfrac{x - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} \sim N(0,1)$

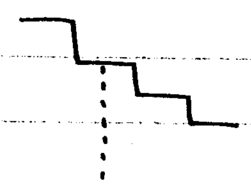
$$\sim \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Can use $N(0,1)$ instead of binomial for reasonably large n (~ 30):

Example 100 flips, what is prob. ≥ 60 heads?

critical value $\rightarrow$ $z = \dfrac{\overset{59.5}{\cancel{60}} - n/2}{\frac{1}{2}\sqrt{n}}$

use 59.5 to be conservative.
get upper bound on area



60 heads

$$= \frac{59.5 - 50}{\frac{1}{2}(10)} = \frac{9.5}{5} = 1.9$$

from table $\displaystyle\int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 0.9713$

1.0000

area above $= \boxed{.0287}$

SIMPLER $\quad z = \dfrac{60 - n/2}{\frac{1}{2}\sqrt{n}} = \dfrac{60-50}{5} = 2$

$$\int_{-\infty}^{z} N(0,1)\,dx = \begin{array}{r} 1.0000 \\ \underline{0.9772} \\ \boxed{.0228} \end{array}$$

1.3.9

Martin Gardner,
How Not to Test a Psychic
Prometheus Books, 1989

$$\frac{x-n}{\frac{1}{2}\sqrt{n}}$$

$$z = \frac{\overset{781}{\cancel{1000}} - 500}{\frac{1}{2}\sqrt{1000}} = \frac{281}{5\cdot\sqrt{10}} = \frac{2.81}{(5)(3.16)} = \frac{281}{15.8} = 17.8!$$

 Pavel Stepanek

presence, and those PS made at home. During the first session, PS made 250 calls in MR's presence, obtaining 162 hits for a score of 162/250 = .64+. Here is one possible scenario:

PS selected four parcels on which to make ten identical calls, and four others on which to make nine identical calls. Later, at home, he adjusted cards in the eight packets so that all 40 + 36 = 76 calls were hits. (Actually, he need adjust only about half of the eight cards—the ones on which he was unlucky.) Of the remaining 250 - 76 = 174, he could expect about half of them, 87, to be hits. Thus his expectation for the session would be 76 + 87 = 163 hits, for a score of 163/250 = .65+. There is no problem about raising his score a comparable amount in the 250 calls made on the same set of parcels at home because he would know the colors of eight cards.

Similar scenarios are easy to devise for the other three sets of packets that would not only explain the percentage of hits but also account for the ways in which ten consecutive calls were split throughout the experiment. Note that if PS selected four packets in each set of 25 parcels for a 10-0 split of calls, and repeated this split on the same four when he recorded calls at home, it would produce 8 × 10 = 40 cases of 10-0 splits. This is just what we are told actually occurred.

Of course there are dozens of different procedures PS could have followed to generate the results given by the published charts, and there is no way we can ever know precisely what happened. However, two conclusions are obvious. By concentrating large numbers of identical calls on a small number of packets, PS could easily have obtained his recorded scores, and of course the procedure would also account for the fantastically strong focusing effect.

As the testing of PS continued, experimenters slowly became aware of the importance of labeling all parts of the test materials as well as imposing better controls—above all of not allowing PS to make calls when he was unsupervised. In the next chapter we shall see how MR and JGP conducted what they considered a more carefully designed test intended to confirm the results of the badly flawed experiment just described.

Before going on to this, however, it is worth mentioning that in May 1962 PS was tested (I do not know where) by H. N. Banerjee, presumably a parapsychologist from India. My information on this rests entirely on a footnote in MR's 1965 monograph (part 3, p. 18). He says he is not including details in his monograph because Banerjee worked alone with PS. The results, he says, were given in *Five Years Report of Seth Sohan Lal Memorial Institute of Parapsychology,* wherever that is, 1963, page 42. I have not tried to run down this report, and can only repeat MR's assertion that out of 1,000 calls on white/green cards PS made 781 hits.

p. 11 of Hossain & Morgan 04 (Do Hossain & Morgan use binomial instead of N(0,1)?)

low reserve (v_*) cases:

NULL HYPOTHESIS: Revenue Equiv.

ONE-SIDED ALTERNATIVE: B outperforms A

Count Treatment B outperforms A: 9 out of 10 for CDs

7 out of 10 for XBOX games

$\boxed{16 \text{ out of } 20}$

"the p-value of the one-sided binomial test is 0.005"

to see this:

$$Z = \frac{x - \frac{n}{2}}{\frac{1}{2}\sqrt{n}} = \frac{16-10}{\frac{1}{2}\sqrt{20}} = \frac{6}{2.23606} = 2.683$$

$$\left(\frac{12}{\sqrt{20}} = \frac{6}{\sqrt{5}} = \frac{6\sqrt{5}}{5}\right)$$

$$\text{table of } \int_{-\infty}^{Z} N(0,1) = 0.99632$$

1.00000 − 0.99632 = .00368 ✓

Normal not good

---

p. 13 CDs in table 5

theory predicts higher revenue under treatment C (vs. A (higher v_*), conditional on being sold!

NULL Hyp.: Rev. equiv.

one-sided: {higher revenue C vs. A | sold}

actually 7/8

$$Z = \frac{\overset{x - n/2}{7-4}}{\frac{1}{2}\sqrt{n}} = \frac{6}{\sqrt{8}} \quad \frac{3}{\sqrt{2}} = 2.12$$

$$p = 1 - .9830 = .027$$

1.0000 − .9830 = .0270 ✓ Normal not good ("0.035")

use binomial

prob. 7/8 + prob 8/8

$= 8/256 + 1/256 = 9/256$

$= .0351$

Bayes' Rule & Hypothesis testing

NULL HYPOTHESIS
(say coin is FAIR.)

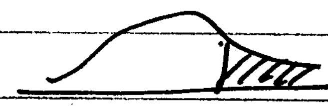
P(DATA | NULL) gives one-tailed statement
about how probable observed
DATA is, given NULL HYPOTHESIS.

TESTING FAIRNESS Example

Suppose we get 65 heads out of 100 flips

UNDER NULL HYP.
$$z = \frac{65 - 50}{\frac{1}{2}\sqrt{100}} = \frac{15}{5} = 3$$

$$\int_{-\infty}^{3} = \begin{array}{r} 1.00000 \\ \underline{0.99865} \\ .00135 \end{array}$$



1-tailed prob. = .00135 → 741:1

Doesn't say anything about P(NULL | DATA)

Bayes' Rule

$$P(\text{NULL}, \text{DATA}) = P(\text{NULL} | \text{DATA}) \cdot P(\text{DATA}) = P(\text{DATA} | \text{NULL}) \cdot P(\text{NULL})$$

$$\boxed{\therefore \quad P(\text{NULL} | \text{DATA}) = \frac{P(\text{DATA} | \text{NULL}) \cdot P(\text{NULL})}{P(\text{DATA})}}$$

Suppose we know (guess?, assume?) $P(\text{NULL})$ (fair coin say)

then

also need

$$P(\text{DATA}) = P(\text{DATA} | \text{NULL}) \cdot P(\text{NULL}) + \boxed{P(\text{DATA} | \overline{\text{NULL}})} \cdot P(\overline{\text{NULL}})$$

so $P(\text{NULL} | \text{DATA})$ is now known in terms of $P(\text{DATA} | \text{NULL})$

Example of ~~Bayesian~~ analysis ~~[illegible]~~ TESTING FOR DISEASE

TEST is + or −

DISEASE is D or Healthy

Let $\boxed{\begin{array}{l} P(+|D) = 0.95 \\ P(+|H) = 0.10 \end{array}}$ $\begin{array}{l} \Rightarrow P(-|D) = 0.05 \\ \Rightarrow P(-|H) = 0.90 \quad \text{(false alarm)} \end{array}$

Assume prior for Bayes' Rule, $P(D) = 0.01$ (rare disease)

prob. of disease if test positive:

$$\boxed{P(D|+) = \frac{P(+|D)\,P(D)}{P(+)}} = \frac{P(+|D)\,P(D)}{P(+|D) \cdot P(D) + \boxed{P(+|H)} \cdot P(H)}$$

$$= \frac{(0.95)(0.01)}{(0.95)(0.01) + \boxed{0.10}(0.99)} = \frac{0.0095}{0.0095 + 0.099}$$

$$= \frac{0.0095}{0.1085} = 0.0876 = 8.76\%$$ chance of having disease if test is +

---

intuition out of 100 tested, 10 healthy will test +

~ 1 Diseased will test +

so about $1/11 \approx 9\%$ of + Tests will actually be sick.