Unit 8: The Normal Distribution

◆□▶ ◆□▶ ◆臣▶ ◆臣▶ 臣 の�?

Ex 1: P(0) = .9, P(1) = .1

n = 1, 3, 10, 25, 400

After *n* draws, the possible sums are 0 to *n*, but the most proba- $^{0.1}$ ble sums group close to .1*n*.





▲ロト ▲圖ト ▲ヨト ▲ヨト 三ヨー のへで

Nine zeros, one one, n draws: number of ones:

Same initial distribution as last slide, but with horizontal scale adjusted so that the possible range of outcomes, 0 to n, is always the same width.

The next slide is similar, with two other initial distributions, both from the text.



(日)、

= √Q (~

One one, one two, one nine, n draws. Sum:

0.03 n = 25 0.08 n = 100 0.025 0.02 0.05 0.015 0.04 0.01 0.005 0.02 50 100 200 20 40 60 80 100 0.0175 n = 50 0.015 0.04 n = 400 0.0125 0.03 0.01 0.0075 0.02 0.005 0.0025 0.01 100 200 300 400 100 200 300 400 0.01 n = 100 0.025 n = 900 0.008 0.02 0.005 0.015 0.004 0.01 0.002 0.005 400 500 800 200 400 600 800

Flipping a coin, number of heads:

▲□▶ ▲圖▶ ▲臣▶ ▲臣▶ = 臣 = のへで

Ex: Law of Averages + Central Limit Theorem

Roll a die *n* times, with n = 6, 12, 24, 48, 96. What are the probabilities of getting exactly *k* 4's?

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 のへぐ

On edstat-1 usenet discussion group:

From: IN%"wvenable@attunga.stats.adelaide.edu.au" 20-JAN-1995
08:59:13.14

Subj: RE: Internet-Based Survey Research

If you mean "surveys" of the kind where someone puts an article in several newsgroups asking people to fill out an electronic questionnaire and return it, the answer has always seemed to me to be simple: Such surveys are ALWAYS a waste of time and resources since they can only yield non-random, indeed self-selected samples from an ill-defined population. The procedure is on exactly the same footing as TV telephone polls, newspaper advertisement polls and the like. The procedure does not to stand up to the most basic scientific criteria of validity and thus the exercise is futile for any formal academic purpose (except, perhaps, journalism). They should be outlawed and attract very serious penalties on the grounds that they are a very blatant abuse of the internet. No offence intended ... Bill Venables, Department of Statistics, The University of

Adelaide, South AUSTRALIA. 5005.

Sample survey: Randomize!

Simple random sample!

- every sample of given size has same chance of being chosen
- theory is based on randomness

Cluster random sample?

- pick some "clusters" out of population and poll everyone in them
- small samples (and there are fewer clusters than individuals) means less randomness

Stratified random sample?

 divide entire population into "strata" and take random sample of each stratum

 if strata are not of equal size, small ones may be over-represented in sample