1 Are all head sizes alike?

Stephen Jay Gould's book "The Mismeasure of Man" discusses a table from a 1978 article by Epstein. Gould read the original article and found that "a glance at E A. Hooton's original table, reproduced below,[1] reveals that the *SE* column had been copied and re-labelled *SD*" Then, using this SD, and the n, to compute a much smaller-than-it-should-be SE, Epstein was able to "show" that the CI's for mean head circumference for people of varied vocational statuses did not overlap, and thus that there were "statistically significant" inter-group differences.

Vocational Status	N	Mean (in mm)	"S.D."
Professional	25	569.9	1.9
Semiprofessional	61	566.5	1.5
Clerical	107	566.2	1.1
Trades	194	565.7	0.8
Public service	25	564.1	2.5
Skilled trades	351	562.9	0.6
Personal services	262	562.7	0.7
Laborers	647	560.7	0.3

i. Explain why the "SDs" in the table should not decrease with increasing n, i.e., why the SD from a smaller n is as likely to be greater than the SD from a bigger n 1 as it is to be smaller. If SD's were smaller (some argue larger) in larger samples, then the SD of the diameters of red blood cells should be different for a large adult than a smaller adult!

ii. Also, from what you have seen of hat-sizes, what makes sense as the SD, and thus the CV, for inter-individual headsizes?

2 Births after The Great Blackout of 1966

On November 9, 1965, the electric power went out in New York City, and it stayed out for a day – The Great Blackout. Nine months later, newspapers suggested that New York was experiencing a baby boom. The table shows the number of babies born every day during a twenty-five day period, centered nine months and ten days after The Great Blackout.

Number of births in New York, Monday August 1-Thursday August 25, 1966.

Mon	Tue	Wed	Thu	Fri	Sat	Sun
451	468	429	448	466	377	344
448	438	455	468	462	405	377
451	497	458	429	434	410	351
467	508	432	426			

These numbers average 436. This turns out to be not unusually high for New York. But there is an interesting twist: the 3 Sundays only average 357.

i. How likely is it that the average of three days chosen at random from the table will be 357 or less? What do you infer? Hint: The SD of the 25 numbers in the table is about 40. Formulate the null hypothesis; the normal approximation can be used.

ii. The above question and the following footnote come from the Statistics text by Freedman et al.

   "Apparently, the New York Times sent a reporter around to a few hospitals on Monday August 8, and Tuesday August 9, nine months after the blackout. The hospitals reported that their obstetric wards were busier than usual – apparently because of the general pattern that weekends are slow, Mondays and Tuesdays are busy. These "findings" were published in a front-page article on Wednesday, August 10, 1966, under the headline "Births Up 9 Months After the Blackout." This seems to be the origin of the baby-boom myth."

   Exercise: Suggest a better plan for estimating the impact, if any, of the Blackout on the number of births.

iii. (Still on the subject of births, but now in Qubec). In an effort to bolster sagging birth rate, the Qubec government in its budget of March 1988 implemented a cash bonus of $4,500 to parents who had a third child. Suggest a method of measuring the impact of this incentive scheme – be both precise and concise.

3 Planning ahead

One has to travel a distance of 7500 Km by 4-wheel jeep, over very rough terrain, with no possibility of repairing a tire that becomes ruptured. Suppose one starts with 14 intact tires (the 4, plus 10 spares). It is known that on average, tires rupture at the rate of 1 per 5,000 tire-Kms (the mean interval between punctures is 5,000 tire-Kms). Assume ruptures occur independently

[1] Table VIII-17 "Mean and standard deviation of head circumference for people of varied vocational statuses" , The American Criminal, v. 1, Harvard U. Press, 1939,

of the of tire position or the distance already driven with the tire (i.e., the sources of failure are purely external). Also, ignore the possibility of multiple failures from a single source, e.g. a short bad section of the trail.

Calculate the probability of completing the trip, using the..

i. Poisson Distribution for the *number* of ruptures.

ii. Exact distribution of a sum of distances i.e. of a (fixed) number of '*distance*' random variables.

iii. Central Limit Theorem to approximate the distribution in ii.

iv. Central Limit Theorem to approximate the distribution in i.

v. Random number fns. in `R/SAS` to simulate intervals between ruptures.

4 A random selection?

A colony of laboratory mice consisted of several hundred animals. Their average weight was about 40 grams, with an SD of about 5 grams. As part of an experiment, graduate students were instructed to choose 25 animals haphazardly, without any definite method. The average weight of these 25 sampled animals was 43 grams. Is choosing animals haphazardly the same as drawing them at random? Assess this by calculating the probability, under strict random selection, of obtaining an average of 43 grams or greater.

5 Planning ahead

On the average, conventioneers weigh about 150 pounds; the SD is 25 pounds.

i. If a large elevator for a convention centre is designed to lift a maximum of 15,500 pounds, the chance it will be overloaded by a random group of 100 conventioneers is closest to which of the following: 0.1 of 1%, 2%, 5%, 50%, 95%, 98%, 99.9% ? Explain your reasoning.

ii. The weights of conventioneers are unlikely to have a Gaussian ("Normal") distribution. In the light of this information, are you still comfortable using the Normal distribution for your calculations in part i? Explain carefully. Explain why the 'random' is key to being able to answer part i. and what impact it would have if it is not the case.

6 An unexpected pattern: or is it?

Data collected on the length of time to diagnose and treat breast cancer show that the diagnostic biopsy results was equally likely to be received on any one of the weekdays from Monday to Friday. Consider the results received the first week of October, say Monday October 1 to Friday October 5. Suppose that the women with positive biopsies then had surgery on one of the weekdays of the last full week of October, i.e., Monday October 22 to Friday October 26. Suppose further that the day of the surgery was also equally likely to be any one of these 5 weekdays, and unrelated to which day the biopsy result was received.

i. Derive and plot the probability distribution of the length of the interval (i.e., the number of days) from when the biopsy result was received until the woman had the surgery. Comment on its shape, and why it is this shape, and what would happen if there were several stages, not just 2.

ii. Calculate the mean and standard deviation of this random variable.

7 A snail's pace

A snail (escargot) starts out to climb a very high wall. During the day it moves upwards an average of 22 cm (SD 4 cm); during the night, independently of how well it does during the day, it slips back down an average of 12 cm (SD 3 cm). The forward and backward movements on one day/night are also independent of those on another day/night.

i. After 16 days and 16 nights, how much vertical progress will it have made? Answer in terms of a mean and SD. Note that – contrary to what many students in previous years calculated – the SD of the total progress made is not 80 cm; show that it is in fact 20 cm.

ii. What is the chance that, after 16 days and 16 nights, it will have progressed more than 150 cm?

iii. "Independence was 'given'. Did you have to make strong [and possibly unjustified] distributional assumptions in order to answer part b? Explain carefully.

8 Student's t-distribution - beyond $n = 10$

"Student"'s table was for $z = (\bar{y} - \mu_0)/s$, not the $t = (\bar{y} - \mu_0)/(s/\sqrt{n})$ tabulated and used today [Also, the s in Student's z was obtained by $\div n$, not $\div(n-1)$].

Moreover, his 1908 table only went up to $n = 10$. For $n > 10$ he suggested using $z = (\bar{y} - \mu_0)/(s/\sqrt{n-3})$ and obtaining the (approximate) p-value by using the Normal table to finding the tail area corresponding to this z value.

His first e.g.'s had $n = 10$, 6 and 2, he "conclude(d) with an example which comes beyond the range of the tables, there being eleven experiments."

For this, he uses the approximation $\Delta \sim N(\bar{d}, s/\sqrt{n-3})$ to arrive at the statement that there is a 0.934 probability "that kiln-dried barley seed gives a higher barley yield than non-kiln-dried seed." [i.e. that $\Delta > 0$ – see below]

i. Use today's packages/functions (e.g. the `pt` function in `R` or `tdist` function in `Excel`, or `probt` in `SAS`) to check how accurate his approximation was in this case.[2] Note that he calculated each SD as $\{(1/11) \times \sum(\text{Diff} - \overline{\text{Diff}})^2\}^{1/2}$.

ii. Do likewise with his other 3 p values (notice the typo in the mean difference in crop value in the last column).

	lbs. head corn per acre			price head corn			cwts. straw per acre			value of crop per acre		
Year	NKD	KD	Diff	NKD	KD	Diff	NKD	KD	Diff	NKD	KD	Diff
1899	1903	2009	+106	26.5	26.5	0	19.25	25	+5.75	140.5	152	+11.5
1899	1935	1915	-20	28	26.5	-1.5	22.75	24	+1.25	152.5	145	-7.5
1899	1910	2011	+101	29.5	28.5	-1	23	24	+1	158.5	161	+2.5
1899	2496	2463	-33	30	29	-1	23	28	+5	204.5	199.5	-5
1899	2108	2180	+72	27.5	27	-0.5	22.5	22.5	0	162	164	+2
1899	1961	1925	-36	26	26	0	19.75	19.5	-0.25	142	139.5	-2.5
1899	2060	2122	+62	29	26	-3	24.5	22.25	-2.25	168	155	-13
1900	1444	1482	+38	29.5	28.5	-1	15.5	16	+0.5	118	117.5	-0.5
1900	1612	1542	-70	28.5	28	-0.5	18	17.25	-0.75	128.5	121	-7.5
1900	1316	1443	+127	30	29	-1	14.25	15.75	+1.5	109.5	116.5	+7
1900	1511	1535	+24	28.5	28	-0.5	17	17.25	+0.25	120	120.5	+0.5
Ave.	1841.5	1875.2	+33.7	28.45	27.55	-0.91	19.95	21.05	+1.10	145.82	144.68	+1.14
SD	–	–	63.1	–	–	0.79	–	–	2.25	–	–	6.67
SD/$\sqrt{8}$	–	–	22.3	–	–	0.28	–	–	0.80	–	–	2.40

Excerpts from section IX of Student's 1908 paper... To test whether it is of advantage to kiln-dry barley seed before sowing, seven varieties of barley were sown (both kiln-dried [KD] and not kiln-dried [NKD]) in 1899 and four in 1900; the results are given in the table. (corn price is in shillings per quarter and the value of the crop is in shillings per acre).

It will he noticed that the kiln-dried seed gave on an average the larger yield of corn and straw, but that the quality was almost always inferior. At first sight this might be supposed to be due to superior germinating power in the kiln-dried seed, but my farming friends tell me that the effect of this would be that the kiln-dried seed would produce the better quality barley. Dr Voelcker draws the conclusion: "In such seasons as 1899 and 1900 there is no particular advantage in kiln-drying before mowing." Our examination completely justifies this and adds "and the quality of the resulting barley is inferior though the yield may be greater."

In this case I propose to use the approximation given by the normal curve with standard deviation $s/\sqrt{n-3}$ and therefore use Sheppard's (Normal) tables, looking up the difference divided by $s/\sqrt{8}$. The probability in the case of yield of corn per acre is given by looking up $33.7/22.3 = 1.51$ in Sheppard's tables. **This gives $p = 0.934$**, or the odds are about 14 to 1 that kiln-dried corn gives the higher yield.

Similarly $0.91/0.28 = 3.25$, corresponding to $p = 0.9994$[3] so that the odds are very great that kiln-dried seed gives barley of a worse quality than seed which has not been kiln-dried.

Similarly, it is about 11 to 1 that kiln-dried seed gives more straw and about 2 to 1 that the total value of the crop is less with kiln-dried seed.

[2] Others had to wait for his extended z table published in 1917, in order to obtain the exact probability.

[3] As pointed out in §V, the normal curve gives too large a value for p when the probability is large. I find the true value in this case to be $p = 0.9976$. It matters little, however, to a conclusion of this kind whether the odds in its favour are 1660 to 1 or merely 416 to 1.

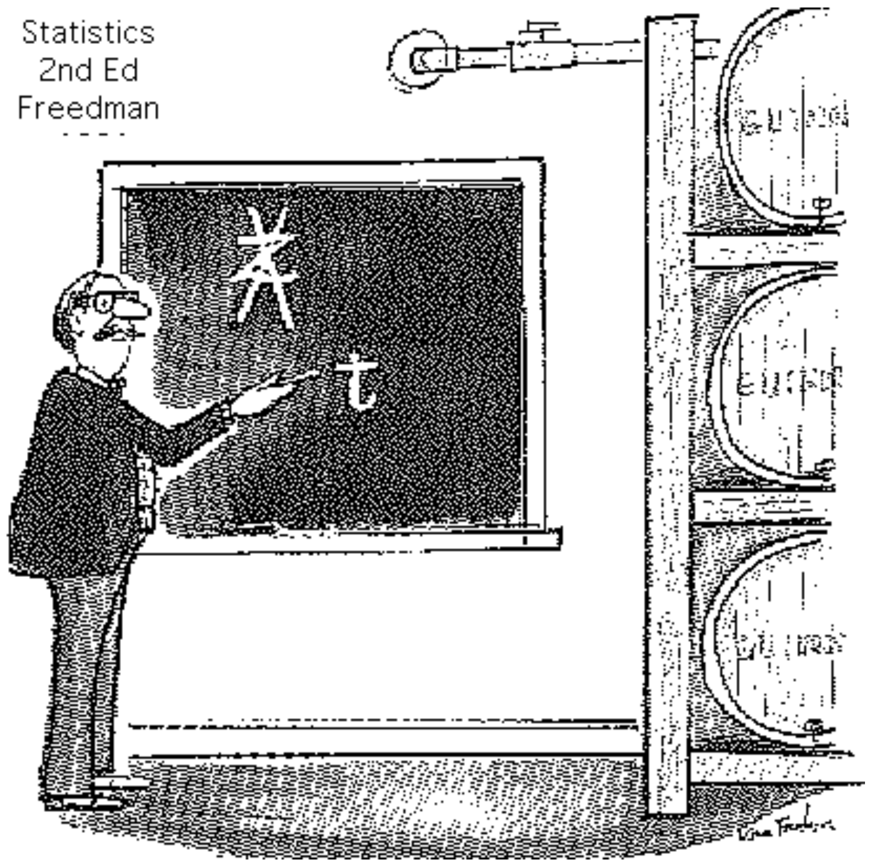

Figure 1: The cartoon, from the textbook *Statistics* by Freedman, Pisani and Purves, refers to switching from the ratio $(\bar{y}-\mu_Y)/(\sigma/\sqrt{n})$ (where σ is known) to the ratio $(\bar{y}-\mu_Y)/(s/\sqrt{n})$ (where s is an estimate of the unknown σ). Ironically, there is another z as well: in 1908 Student derived and tabulated the distribution of the ratio: $z = (\bar{y}-\mu_Y)/s^*$, with s^* obtained using a divisor of n. Later, in the mid 1920s, Fisher got him to switch to the ratio $(\bar{y}-\mu_Y)/(s/\sqrt{n})$, with s obtained using a divisor of $n-1$. It appears that Student was the one who made the name change from $Student's\, z$ to $Student's\, t$, and Fisher who did the heavy math lifting, and who saw the much wider applicability of the t distribution. Fisher saw a t r.v. as (proportional to) the ratio of a Gaussian r.v. to the square root of an independent r.v. with a chi-squared distribution, and the centrality of the concept of 'degrees of freedom'. For more, see 2008 article by JH MJ and EM under Resources.

'Student' in 1908

Figure 2: from `http://www.york.ac.uk/depts/maths/histstat/people/`

9 Experiments to Determine the Density of the Earth. By Henry Cavendish, Esq. F.R.S. and A. S.

The 29 measurements (cf. an earlier assignment sheet) are repeated here:

5.5 5.61 4.88 5.07 5.26 5.55 5.36 5.29 5.58 5.65 5.57 5.53 5.62 5.29 5.44 5.34 5.79 5.1 5.27 5.39 5.42 5.47 5.63 5.34 5.46 5.3 5.75 5.68 5.85

The following is from pp 521-522 of his report.

> From this table it appears, that though the experiments agree pretty well together, yet the difference between them, both in the quantity of motion of the arm and in the time of vibration, is greater than can proceed merely from the error of observation. As to the difference in the motion of the arm, it may very well be accounted for, from the current of air produced by the difference of temperature; but, whether this can accounted for the difference in the time of vibration, is doubtful. If the current of air was regular, and of the same swiftness in all parts of the vibration of the ball, I think it could not; but, as there will most likely be much irregularity in the current, it may very likely be sufficient to account for the difference.
>
> By a mean of the experiments rnade with the wire first used, the density of the earth comes out **5.48** times greater than that of water; and by a mean of those made with tire latter wire, it comes out the same; and the extreme difference of the results of the 23 observations made with this wire, is only **.75**; so that the extreme results do not differ from the mean by more than **.38**, or $\frac{1}{14}$ of the whole, and therefore the density should seen to be deterinined hereby, to great exactness.
>
> It, indeed, may be objected, that as the result appears to be influenced by the current of air, or some other cause, the laws of which we are not well acquainted with, this cause may perhaps act always, or commonly, in the same direction, and thereby make a considerable error in the result. But yet, as the experiments were tried in various weathers, and with considerable variety in the difference of temperature of the weights and air, and with the arm resting at different distances from the sides of the case, it seems very unlikely that this cause should act so uniformly in the same way, as to make the error of the mean result nearly equal to the difference between this and the extreme; and, therefore, it seem very unlikely that the density of the earth should differ from 5.48 by so much as $\frac{1}{14}$ of the whole.
>
> Another objection, perhaps, may be made to these experiments, namely, that it is uncertain whether, in these small distances, the force of gravity follows exactiy the same law as in greater distances. There is no reason, however, to think that any irregularity of this kind takes pIace, until the bodies come within the action of what is called the attraction of cohesion, and which seems to extend only to very minute distances. With a view to see whether the result could be affected by this attraction, I made the 9th, 10th, 11th and 15th experiments, in which the balls were made to rest as clese to the sides of the case as they could; but there is no difference to be depended on, between the results under that circumstance, and when the balls are placed in any other part of the case.
>
> According to the experiments made by Dr. MASKELYNE on the attraction of the hill Schehallien, the density of the earth is $4\frac{1}{2}$ times that of water; which differs rather more from the preceding determination than I should have expected. But I forbear entering into any consideration of which determination is most to be depended on, till 1 have examined more carefully how much the preceding determination is affected by irregularities whose quaantity I cannot measure.

Exercise

i. Find the mean of the 29 values.

ii. Calculate a 95% CI to accompany it.

iii. Would a "trimmed mean" be useful here?
(see `http://en.wikipedia.org/wiki/Truncated_mean`).