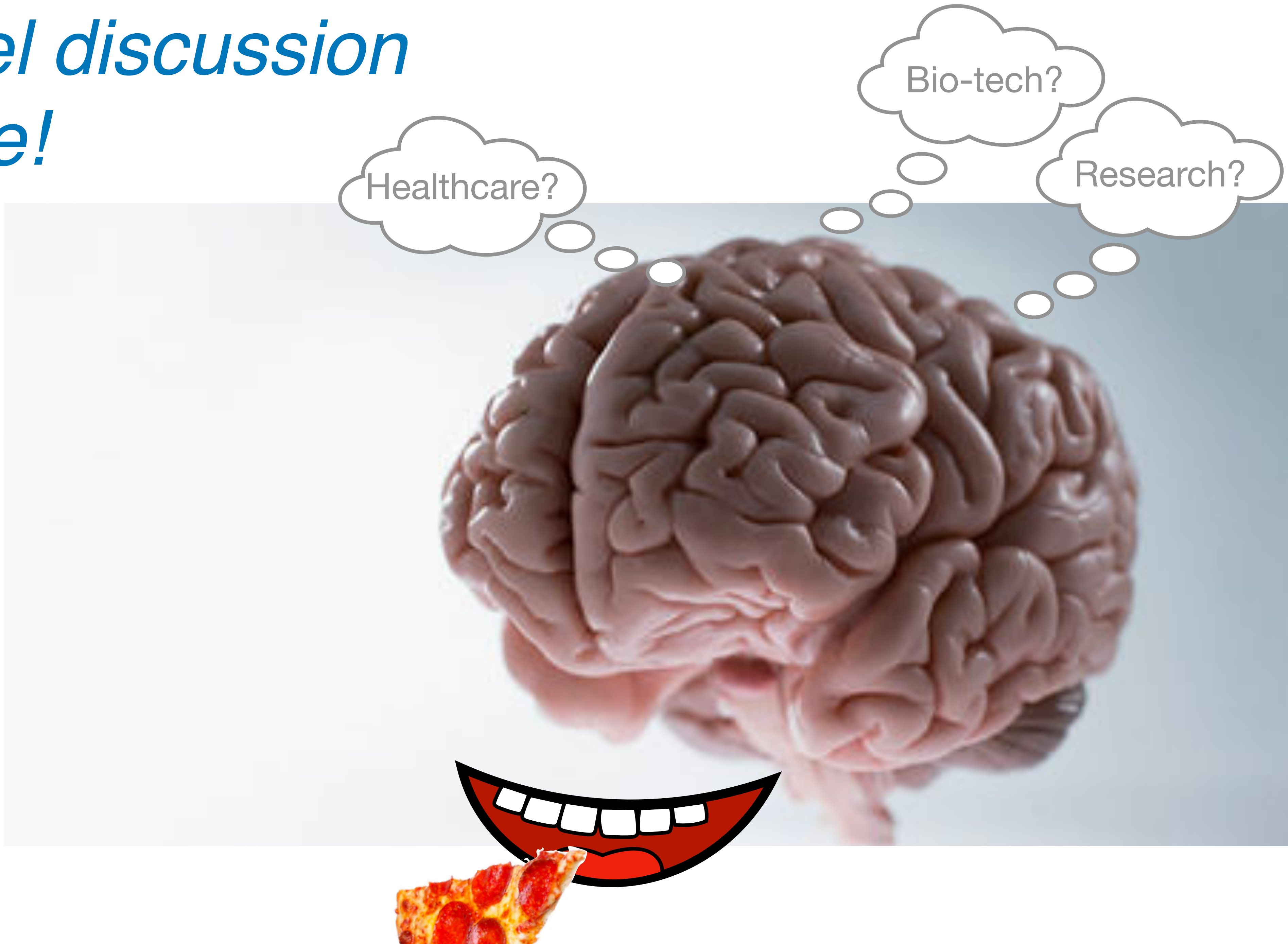
STATISTICS WITH GAUSSIANS

10.22.2018

Department of Neuroscience PIZZA WITH PROFESSORS presents:

Student Organizations…
How can I (you) get involved?

Pizza and panel discussion
All are welcome!
Meet friends!
Get involved!



Monday
October 22
5:00 p.m.
NHB 1.720

RECAP

* central limit theorem!
* gaussian distribution
* gaussian transformations
* estimating mean & variance from sample

SCIPY.STATS.NORM

- **scipy.stats** has many useful functions
- **from scipy.stats import norm**
  - **norm.pdf** - the probability density function
  - **norm.cdf** - the cumulative density function
  - **norm.sf** - “survival function” (1 - cdf)
  - **norm.isf** - inverse survival function
- *all functions assume a standard normal unless you give them extra arguments!*

GAUSSIAN STANDARD ERROR

* if we have samples from a Gaussian we can always use bootstrapping to find the standard error

* but, like the binomial distribution, there is also an analytic solution for the standard error:

$$SE = \frac{s}{\sqrt{n}}$$

GAUSSIAN CONFIDENCE INTERVAL

* suppose we have X~N(0,1), a standard normal RV

* what’s its 95% confidence interval? (i.e. what is the interval [a,b] where 95% of the samples from X will fall between a and b?)

GAUSSIAN CONFIDENCE INTERVAL

* for [a,b] we want a to be the 2.5th %-ile and b to be the 97.5th %-ile

* so:

GAUSSIAN CONFIDENCE INTERVAL

* for [a,b] we want a to be the 2.5th %-ile and b to be the 97.5th %-ile

* so: a=norm.isf(0.975), b=norm.isf(0.025)

GAUSSIAN CONFIDENCE INTERVAL

* for a standard normal, the 95% confidence interval is almost exactly [-1.96, 1.96]

GAUSSIAN CONFIDENCE INTERVAL

- suppose we have a sample from a gaussian RV with sample mean **μ** and sample variance **s²**, what's the 95% confidence interval for the true mean?

-

GAUSSIAN CONFIDENCE INTERVAL

* suppose we have a sample from a gaussian RV with sample mean **μ** and sample variance **s^2**, what's the 95% confidence interval for the true mean?

* [μ - 1.96*SE, μ + 1.96*SE]

GAUSSIAN Z-TEST

* recall the definition of the p-value: the probability of observing a result *at least as extreme* as your actual result

* to test whether a particular value is unlikely under a gaussian distribution, we can evaluate the CDF of that distribution!

GAUSSIAN Z-TEST

* e.g. we know that a specific voxel recorded with fMRI has a background response of μ=0.5 and σ^2=1.7

* we record a response of x=2.4. is this significantly <u>higher</u> than the background response?

GAUSSIAN Z-TEST

* first we compute a “z-score” for x:

  * $z(x) = (x - \mu) / \sigma$

* then we can compute Pr[X>=x] using **norm.sf**

END