GAUSSIANS

10.7.2020

RECAP

* bootstrap
  * confidence interval
  * standard error (standard deviation of bootstrapped statistic)

CENTRAL LIMIT THEOREM

* DEMO: let’s take a bunch of random values (from the same distribution)

* and take their average

* let’s do this many times

CENTRAL LIMIT THEOREM

* as the number of samples we take increases, the distribution of their averages converges to..

* a Gaussian distribution!

CENTRAL LIMIT THEOREM

* and *it doesn’t matter what distribution you start with*

* *THE DISTRIBUTION YOU START WITH CAN BE TOTALLY BIZARRE AND NOTHING LIKE A GAUSSIAN DISTRIBUTION*

* ***AND IT STILL WORKS!!!!!!***

CENTRAL LIMIT THEOREM

* this is why gaussian distributions are everywhere (fMRI data, behavioral data, calcium imaging data, etc. etc. etc.), and we will learn about them

GAUSSIANS

Carl Gauss

* "Gaussian distribution"
  aka
  "Normal distribution"

* X ~ N(μ,σ²) : X is a normal RV with mean μ and variance σ²

$$PDF(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

GAUSSIANS

* “Standard normal” distribution has mean $\mu=0$, and variance $\sigma^2=1$

$$PDF \propto e^{-x^2}$$

* **np.random.randn** generates random variables from standard normal

GAUSSIAN TRANSFORMATIONS

* suppose you have samples from $X \sim N(0,1)$ but you want samples from $Y \sim N(5,1)$

* i.e. you want samples from a distribution with $\mu=5$

*

GAUSSIAN TRANSFORMATIONS

* suppose you have samples from $X \sim N(0,1)$ but you want samples from $Y \sim N(5,1)$

* i.e. you want samples from a distribution with $\mu=5$

* add the mean! $(X + b) \sim N(b,1)$

GAUSSIAN TRANSFORMATIONS

```
* suppose you have samples from X ~ N(0,1)
  but you want samples from Y ~ N(0,5)

* i.e. you want samples with variance 5

* multiply by sqrt(5)!

  (a * X) ~ N(0,sqrt(a))
```

ESTIMATING MEAN & VARIANCE

* “sample mean” is just the mean of your sample

* but “sample variance” (s^2) is *not* just the variance of your sample! (omg)

  * i.e. $s^2 \mathrel{!=} \sigma^2$

ESTIMATING MEAN & VARIANCE

* sample variance (s^2) is defined as:

$$s^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2$$

*

ESTIMATING MEAN & VARIANCE

* sample variance (s^2) is defined as:

$$s^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2$$

* which is normalized by n-1 instead of n

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:

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

END