CS70: Lecture 27

1. Review: Continuous Probability
2. Bayes' Rule with Continuous RVs
3. Normal Distribution
4. Central Limit Theorem
5. Confidence Intervals
6. Wrapup.

Continuous Probability

1. pdf: $Pr[X \in (x, x+\delta]] = f_X(x)\delta$.
2. CDF: $Pr[X \le x] = F_X(x) = \int_{-\infty}^{x} f_X(y)dy$.
3. $U[a,b]$, $Expo(\lambda)$, target.
4. Expectation: $E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$.
5. Variance: $var[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$.
6. Variance of Sum of Independent RVs: If X_n are pairwise independent, $var[X_1 + \cdots + X_n] = var[X_1] + \cdots + var[X_n]$

Continuous RV and Bayes' Rule

Example 1:

W.p. $1/2$, X, Y are i.i.d. $Expo(1)$ and w.p. $1/2$, they are i.i.d. $Expo(3)$.

Calculate $E[Y|X = x]$.

Let B be the event that $X \in [x, x+\delta]$ where $0 < \delta \ll 1$.

Let A be the event that X, Y are $Expo(1)$.

Then,

$$\begin{aligned} Pr[A|B] &= \frac{(1/2)Pr[B|A]}{(1/2)Pr[B|A] + (1/2)Pr[B|\bar{A}]} = \frac{\exp\{-x\}\delta}{\exp\{-x\}\delta + 3\exp\{-3x\}\delta} \\ &= \frac{\exp\{-x\}}{\exp\{-x\} + 3\exp\{-3x\}} = \frac{e^{2x}}{3 + e^{2x}}. \end{aligned}$$

Now,

$$\begin{aligned} E[Y|X = x] &= E[Y|A]Pr[A|X = x] + E[Y|\bar{A}]Pr[\bar{A}|X = x] \\ &= 1 \times Pr[A|X = x] + (1/3)Pr[\bar{A}|X = x] \ldots = \frac{1 + e^{2x}}{3 + e^{2x}}. \end{aligned}$$

We used $Pr[Z \in [x, x+\delta]] \approx f_Z(x)\delta$ and given A one has $f_X(x) = \exp\{-x\}$ whereas given $\bar{A}$ one has $f_X(x) = 3\exp\{-3x\}$.

Continuous RV and Bayes' Rule

Example 2:

W.p. $1/2$, Bob is a good dart player and shoots uniformly in a circle with radius 1. Otherwise, Bob is a very good dart player and shoots uniformly in a circle with radius $1/2$.

The first dart of Bob is at distance 0.3 from the center of the target.

(a) What is the probability that he is a very good dart player?

(b) What is the expected distance of his second dart to the center of the target?

Note: If uniform in radius r, then $Pr[X \le x] = (\pi x^2)/(\pi r^2)$, so that $f_X(x) = 2x/(r^2)$.
(a) We use Bayes' Rule:

$$\begin{aligned} Pr[VG|0.3] &= \frac{Pr[VG]Pr[\approx 0.3|VG]}{Pr[VG]Pr[\approx 0.3|VG] + Pr[G]Pr[\approx 0.3|G]} \\ &= \frac{0.5 \times 2(0.3^2)\varepsilon/(0.5^2)}{0.5 \times 2(0.3^2)\varepsilon/(0.5^2) + 0.5 \times 2\varepsilon(0.3^2)} = 0.8. \end{aligned}$$

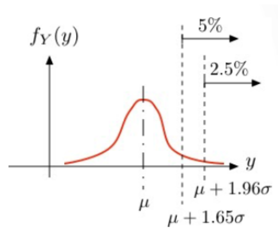
(b) $E[X] = 0.8 \times 0.5 \times \frac{2}{3} + 0.2 \times \frac{2}{3} = 0.4$.

Normal (Gaussian) Distribution.

For any μ and σ, a **normal** (aka **Gaussian**) random variable Y, which we write as $Y = \mathcal{N}(\mu, \sigma^2)$, has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}.$$

Standard normal has $\mu = 0$ and $\sigma = 1$.



Note: $Pr[|Y - \mu| > 1.65\sigma] = 10\%$; $Pr[|Y - \mu| > 2\sigma] = 5\%$.

Scaling and Shifting and properties

Theorem Let $X = \mathcal{N}(0,1)$ and $Y = \mu + \sigma X$. Then

$$Y = \mathcal{N}(\mu, \sigma^2).$$

Theorem If $Y = \mathcal{N}(\mu, \sigma^2)$, then

$$E[Y] = \mu \text{ and } var[Y] = \sigma^2.$$

Review: Law of Large Numbers.

Theorem: Set of independent identically distributed random variables, X_i,
$A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expectation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is σ^2/n.

Used Chebyshev.

$$Pr[|A_n - \mu| > \varepsilon] \le \frac{var[A_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon} \to 0.$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \le \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Note:

$$E(S_n) = \frac{1}{\sigma/\sqrt{n}}(E(A_n) - \mu) = 0$$

$$Var(S_n) = \frac{1}{\sigma^2/n} Var(A_n) = 1.$$

CI for Mean

Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$

The CLT states that

$$\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \to \mathcal{N}(0,1) \text{ as } n \to \infty.$$

Also,

$$[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] \text{ is a 95\%} - \text{CI for } \mu.$$

Recall: Using Chebyshev, we found that

$$[A_n - 4.5\frac{\sigma}{\sqrt{n}}, A_n + 4.5\frac{\sigma}{\sqrt{n}}] \text{ is a 95\%} - \text{CI for } \mu.$$

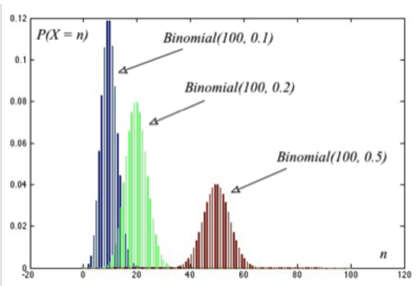
Thus, the CLT provides a smaller confidence interval.

Coins and normal.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$.

Here, $\mu = p$ and $\sigma = \sqrt{p(1-p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1).$$



Coins and normal.

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$.
Here, $\mu = p$ and $\sigma = \sqrt{p(1-p)}$. CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1)$$

and

$$[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] \text{ is a 95\%} - \text{CI for } \mu$$

with $A_n = (X_1 + \cdots + X_n)/n$.
Hence,

$$[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] \text{ is a 95\%} - \text{CI for } p.$$

Since $\sigma \le 0.5$,

$$[A_n - 2\frac{0.5}{\sqrt{n}}, A_n + 2\frac{0.5}{\sqrt{n}}] \text{ is a 95\%} - \text{CI for } p.$$

Thus,

$$[A_n - \frac{1}{\sqrt{n}}, A_n + \frac{1}{\sqrt{n}}] \text{ is a 95\%} - \text{CI for } p.$$

Application: Polling.

How many people should one poll to estimate the fraction of votes that will go for Trump?

Say we want to estimate that fraction within 3% (margin of error), with 95% confidence.

This means that if the fraction is p, we want an estimate $\hat{p}$ such that

$$Pr[\hat{p} - 0.03 < p < \hat{p} + 0.03] \ge 95\%.$$

We choose $\hat{p} = \frac{X_1 + \cdots + X_n}{n}$ where $X_m = 1$ if person m says she will vote for Trump, 0 otherwise.

We assume X_m are i.i.d. $B(p)$.

Thus, $\hat{p} \pm \frac{1}{\sqrt{n}}$ is a 95%-confidence interval for p. We need

$$\frac{1}{\sqrt{n}} = 0.03, \text{ i.e., } n = 1112.$$

Summary

1. Bayes' Rule: Replace $\{X = x\}$ by $\{X \in (x, x+\varepsilon)\}$.
2. Gaussian: $\mathcal{N}(\mu, \sigma^2) : f_X(x) = ...$ "bell curve"
3. CLT: X_n i.i.d. $\implies \frac{A_n - \mu}{\sigma/\sqrt{n}} \to \mathcal{N}(0,1)$
4. CI: $[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] =$ 95%-CI for μ.

CS70: Wrapping Up.

Random Thoughts

Confusing Statistics: Simpson's Paradox

College	F. Appl.	F. Adm.	% F. Adm.	M. Appl.	M. Adm.	% M. Adm.
A	980	490	50%	200	80	40%
B	20	20	100%	800	720	90%
Total	1000	510	51%	1000	800	80%

Applications/admissions of males and females to two colleges of a university.

Male admission rate 80% but female 51%!

However, the admission rate is larger for female students in both colleges....

Female students apply more to the college that admits fewer students.

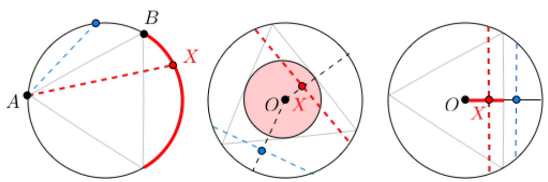
Side note: average high school GPA is higher for female students.

More on Confusing Statistics

Statistics are often confusing:

- The average household annual income in the US is $72k. Yes, the median is $52k.
- The false alarm rate for prostate cancer is only 1%.
  Still only 1 person in 8,000 has that cancer. Prior.
  $\implies$ there are 80 false alarms for each actual case.
- The Texas sharpshooter fallacy:
  Shoot a barn. Paint target cluster. I am sharpshooter!
  People living close to power lines.
  You find clusters of cancers!
  Also find such clusters when looking at people eating kale!
- False causation. Vaccines cause autism.
  Both vaccination and autism rates increased....
- Beware of statistics reported in the media!

Choosing at Random: Bertrand's Paradox



The figures corresponds to three ways of choosing a chord "at random."
Probability chord is larger than $|AB|$ of an inscribed equilateral triangle?

- Choose a point A, choose second point X uniformly on circumference (left): $1/3$
- Choose a point X uniformly in the circle and draw chord perpendicular to the radius that goes through X (center): $1/4$
- Choose a point X uniformly on a given radius and draw the chord perpendicular to the radius that goes through X (right): $1/2$

Confirmation Bias

Confirmation bias: tendency to search for, interpret, and recall information in a way that confirms one's beliefs or hypotheses, while giving less consideration to alternative possibilities.

Confirmation biases contribute to overconfidence in personal beliefs and can maintain or strengthen beliefs in the face of contrary evidence.

Three aspects:

- Biased search for information.
  E.g., facebook friends effect, ignoring inconvenient articles.
- Biased interpretation.
  E.g., valuing confirming versus contrary evidence.
- Biased memory.
  E.g., remember facts that confirm beliefs and forget others.

Confirmation Bias: An experiment

There are two bags.

One with 60% red balls and 40% blue balls;
the other with the opposite fractions.

One selects one of the two bags.

As one draws balls one at time, one asks people to declare whether they think one draws from the first or second bag.

Surprisingly, people tend to be reinforced in their original belief, even when the evidence accumulates against it.

Report Data not Opinion!

A bag with 60% red, 40% blue or vice versa.

Each person pulls ball, reports opinion on which bag:
Says "majority blue" or "majority red."

Does not say what color their ball is.

What happens if first two get blue balls?

Third hears two blue, so says blue, whatever she sees.
Plus Induction.

Everyone says blue...forever ...and ever.

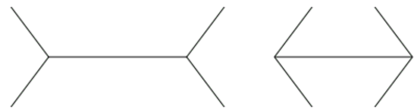
Problem: Each person reported honest opinion rather than data!

Being Rational: 'Thinking, Fast and Slow'

In this book, Daniel Kahneman discusses examples of our irrationality.

Here are a few examples:

- A judge rolls a die in the morning.
  In the afternoon, he has to sentence a criminal.
  Statistically, morning roll high $\Longrightarrow$ sentence is high.
- People tend to be more convinced by articles printed in Times Roman instead of Computer Modern Sans Serif.
- Perception illusions: Which horizontal line is longer?



It is difficult to think clearly!

What to Remember?

Professor, what should I remember about probability from this course?
I mean, after the final.

Here is what the prof. remembers:

- Given the uncertainty around us, understand some probability.
- One key idea - what we learn from observations: the role of the prior; Bayes' rule; Estimation; confidence intervals... quantifying our degree of certainty.
- This clear thinking invites us to question vague statements, and to convert them into precise ideas.

What's Next?

Professor, I loved this course so much!
I want to learn more about discrete math and probability!

Funny you should ask! How about

- CS170: Efficient Algorithms and Intractable Problems a.k.a. Introduction to CS Theory: Graphs, Dynamic Programming, Complexity.
- EE126: Probability in EECS: An Application-Driven Course: PageRank, Digital Links, Tracking, Speech Recognition, Planning, etc.
  Hands on labs with python experiments (GPS, Shazam, ...).
- CS188: Artificial Intelligence: Hidden Markov Chains, Bayes Networks, Neural Networks.
- CS189: Introduction to Machine Learning: Regression, Neural Networks, Learning, etc. Programming experiments with real-world applications.
- EE121: Digital Communication: Coding for communication and storage.
- EE223: Stochastic Control.
- EE229A: Information Theory; EE229B: Coding Theory.

Final Thoughts

More precisely: Some thoughts about the final ....

How to study for the final?

- Lecture Slides; Notes; Discussion Problems; HW
- TA Office Hours, Prof. Office Hours, Reviews by TAs
- Next week: reviews during normal lecture hours:
  - Concept Review (Tuesday);
  - Question Review (Thursday).

Parting Thoughts

You have learned a lot in this course!

Proofs, Graphs, Mod(p), RSA, Reed-Solomon, Decidability, Probability, ... ,
how to handle stress, how to sleep less, how to keep smiling, ...
Difficult course? Yes! Mind expanding? I hope
Useful? You bet!
Finally,

Thanks for taking the course!

Thanks to the CS70 Staff:

- The Terrific Tutors
- The Rigorous Readers
- The Thrilling TAs
- The Amazing Assistants

See you on Tuesday.