CS70: Jean Walrand: Lecture 37.

Gaussian RVs and CLT

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

Continuous Probability

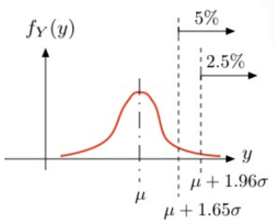
1. pdf: $Pr[X \in (x, x+\delta]] = f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = \int_{-\infty}^{x} f_X(y)dy$.
3. $U[a,b]$, $Expo(\lambda)$, target.
4. Expectation: $E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$.
5. Expectation of function: $E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x)dx$.
6. Variance: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

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

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

$$\begin{aligned} f_Y(y) &= \frac{1}{\sigma} f_X\left(\frac{y-\mu}{\sigma}\right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{(y-\mu)^2}{2\sigma^2}\}. \quad \square \end{aligned}$$

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathcal{N}(0,1)$ since $Y = \mu + \sigma X, \ldots$.

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

First note that $E[X] = 0$, by symmetry.

$$\begin{aligned} var[X] &= E[X^2] = \int x^2 \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\} dx \\ &= -\frac{1}{\sqrt{2\pi}} \int x d\exp\{-\frac{x^2}{2}\} = \frac{1}{\sqrt{2\pi}} \int \exp\{-\frac{x^2}{2}\} dx \text{ by IBP}^1 \\ &= \int f_X(x)dx = 1. \quad \square \end{aligned}$$

[1] Integration by Parts: $\int_a^b f dg = [fg]_a^b - \int_a^b g df$.

Review: Law of Large Numbers.

Theorem: For any 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.

Thus,

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

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,

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

That is,

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

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

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

Thus, for $n \gg 1$, one has

$$Pr[-2 \leq |\frac{A_n - \mu}{\sigma/\sqrt{n}}| \leq 2] \approx 95\%.$$

Equivalently,

$$Pr[\mu \in [A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]] \approx 95\%.$$

That is,

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

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

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

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

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

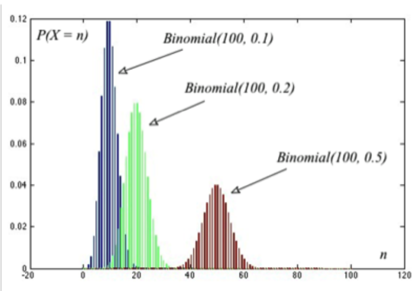
Thus, the CLT provides a smaller confidence interval.

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

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



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

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

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 $\sigma \leq 0.5$,

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

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

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

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

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

Application: Testing Lightbulbs.

Assume that lightbulbs have i.i.d. $Expo(\lambda)$ lifetimes. We want to make sure that $\lambda^{-1} > 1$. Say that we measure the average lifetime A_n of $n = 100$ bulbs and we find that it is equal to 1.2.

What is the confidence that we have that $\lambda^{-1} > 1$? We have,

$$\frac{A_n - \lambda^{-1}}{\lambda^{-1}/\sqrt{n}} = \sqrt{n}(\lambda A_n - 1) \approx \mathcal{N}(0,1).$$

Thus,

$$Pr[\sqrt{n}(\lambda A_n - 1) > \sqrt{n}(\lambda 1.2 - 1)] \approx Pr[\mathcal{N}(0,1) > \sqrt{n}(\lambda 1.2 - 1)].$$

If $\lambda^{-1} < 1$, this probability is at most
$Pr[\mathcal{N}(0,1) > \sqrt{n}(1.2-1)] = Pr[\mathcal{N}(0,1) > 2] = 2.5\%$.

Thus, we conclude that $Pr[\lambda^{-1} > 1] \geq 97.5\%$.

Summary

Gaussian and CLT

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

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,

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

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 $Pr[X \leq x] = (\pi x^2)/(\pi r^2)$, 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) $E[X] = 0.8 \times 0.5 \times \frac{2}{3} + 0.2 \times \frac{2}{3} = 0.4$.