CS70 - Lecture 27 Notes

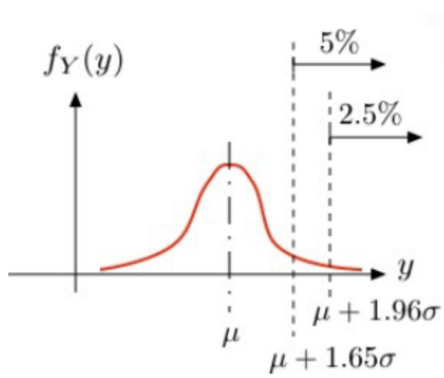
Name: Felix Su SID: 25794773

Spring 2016 GSI: Gerald Zhang

Review: Continuous Probability

1. **pdf**: $\Pr[X \in (x, x+\delta]] = f_X(x)\delta$.
   - Probability of point in $\Omega = 0$, so define prob. of events as intervals = pdf(δ)
   - Pdf is non-negative and integrates to 1
2. **CDF**: $\Pr[X \leq x] = F_X(x) = \int_{-\infty}^{x} f_X(y)dy$.
   - $\Pr[a < x \leq b] = \Pr[X \leq b] - \Pr[X \leq a]$
3. **U[a,b], Expo(λ), target**
4. **Expectation**: $E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$.
   - $x*$probability of $X = x$ in that interval
5. **Expectation of function**: $E[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x)dx$.
6. **Variance**: $\text{Var}[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$.
7. **Gaussian**: $N(\mu, \sigma^2) : f_X(x) = \dots$ "bell curve"
   - When you add up many small RVs, the CDF comes out as a bell shape

Normal Distribution



- For any mean: μ and std. dev: 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: $\Pr[|Y - \mu| > 1.65\sigma] = 10\%$;$\Pr[|Y - \mu| > 2\sigma] = 5\%$.
- Guassian RV is within 2σ of the mean with 95%

Normal Distribution
For any μ and σ, a Gaussian RV, $Y = \mathcal{N}(\mu, \sigma^2)$ has pdf:

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

Scaling and Shifting

> 1. If you scale a Gaussian RV (by $Y = \mu + \sigma X$), you get another Gaussian RV
> 2. When you multiply at RV by a constant (σ), you multiply its variance by the square of that constant (σ^2)

- **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, $f_Y(y) = \frac{1}{\sigma} f_X(\frac{y-\mu}{\sigma})$ (See Lec. 26, slide 19.)
  - $= \frac{1}{\sqrt{2\pi\sigma 2}} \exp\{-\frac{(y-\mu)^2}{2\sigma^2}\}$

Expectation, Variance

- **Theorem** If $Y = \mathcal{N}(\mu, \sigma^2)$, then
  - $E[Y] = \mu$ and $\text{Var}[Y] = \sigma^2$
- **Proof:** It suffices to show the result for $X = \mathcal{N}(0, 1)$ since $Y = \mu + \sigma X, \ldots$
- Thus, $f_X(x) = \frac{1}{2\pi} \exp\{-\frac{x^2}{2}\}$.
  - First note that $E[X] = 0$, by symmetry.
  - $\text{Var}[X] = E[X^2] = \int x^2 \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\} dx$
  - $= -\frac{1}{\sqrt{2\pi}} \int x d \exp\{-\frac{x^2}{2}\}$
  - $= \frac{1}{\sqrt{2\pi}} \int \exp\{-\frac{x^2}{2}\} dx$ (Integration by Parts: $\int_a^b f dg = [fg]_a^b - \int_a^b g df$)
  - $= \int f_X(x) dx = 1$

Central Limit Theorem

> 1. Tells us how many samples to take in order for the airthmetic mean to tend to the expected value of an RV
> 2. Derive the Normalized Sample mean $S_n = \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$
> 3. **CLT:** As $n \to \infty$, the distribution of $S_n \to$ the standard normal distribution $\mathcal{N}(0, 1)$
>   - Expectation of S_n is always 0
>   - Variance of S_n is always 1

- **Law of Large Numbers:** 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.

- Let $S_n = \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$
- Then, $E[S_n] = \frac{1}{\sigma/\sqrt{n}}(E[An] - \mu) = 0$
- $\text{Var}[S_n] = \frac{1}{\sigma^2/n}\text{Var}[A_n] = 1$.

- **Central limit theorem:** As n goes to infinity the distribution of S_n approaches the standard normal distribution.
  - Expectation of S_n is always 0
  - Variance of S_n is always 1

Central Limit Theorem

Normalized Sample Mean

$$S_n = \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \tag{2}$$

Expected Value of Normalized Sample Mean

$$E[S_n] = \frac{1}{\sigma/\sqrt{n}}(E[An] - \mu) = 0 \tag{3}$$

Variance of Normalized Sample Mean

$$\text{Var}[S_n] = \frac{1}{\sigma^2/n}\text{Var}[A_n] = 1 \tag{4}$$

Central Limit Theorem

- Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $\text{Var}[X_1] = \sigma^2$.
- Define $S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$.
  - Then, $S_n \to \mathcal{N}(0,1)$,as $n \to \infty$.
  - That is, $\Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx$.
  - **Proof:** See EE126.
- THe CDF of the RV S_n approches the CDF of $\mathcal{N}(0,1)$
  - PDF begins to look like a bell shape
  - CDF looks like the integral of a bell shape

Central Limit Theorem:

$$\Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx \tag{5}$$

Confidence Interval (CI) for Mean

- Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2 and $A_n = \frac{X_1 + \cdots + X_n}{n}$.
- The CLT states that $\frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \to \mathcal{N}(0,1)$ as $n \to \infty$.
  - Thus, for $n >> 1$, one has $\Pr[|\frac{A_n - \mu}{\sigma/\sqrt{n}}| \leq 2] \approx 95\%$.
  - Equivalently, $\Pr[\mu \in [A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]] \approx 95\%$.
  - That is, $[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]$ is a 95% -CI for μ.

Confidence Interval (CI) for Mean

$$\Pr[|\frac{A_n - \mu}{\sigma/\sqrt{n}}| \leq 2] \approx 95\% \tag{6}$$

$$[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] \text{ is a 95\% -CI for } \mu \tag{7}$$

- Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2 and $A_n = \frac{X_1 + \cdots + X_n}{n}$.
- The CLT states that $\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \to \mathcal{N}(0, 1)$ as $n \to \infty$.
  - Also, $[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]$ is a 95% -CI for μ.
  - Recall: Using Chebyshev, we found that (see Lec. 22, slide 6)
  - $[A_n - 4.5\frac{\sigma}{\sqrt{n}}, A_n + 4.5\frac{\sigma}{\sqrt{n}}]$ is a 95% -CI for μ.
- Thus, the CLT provides a smaller confidence interval.
  - Chebyshev works for all values of n.
  - The CLT assumes n is large enough.

Example Question: Question like this will be on the FINAL

Example:
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 $[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]$ is a 95% -CI for μ with $A_n = (X_1 + \cdots + X_n)/n$.
> Hence, $[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]$ is a 95% -CI for p.
> Solve $\frac{d\mathrm{Var}[X]}{dp} = \frac{d\sigma^2}{dp}$, test each result to find the p that returns the max of the Variance
> Substitute the returned p back in to solve for the max value of σ
> Since $\sigma \leq 0.5$, Substitute σ with the upper bound: $[A_n - 2\frac{0.5}{\sqrt{n}}, A_n + 2\frac{0.5}{\sqrt{n}}]$ is a 95% -CI for p.
> Thus, $[A_n - \frac{1}{\sqrt{n}}, A_n + \frac{1}{\sqrt{n}}]$ is a 95% -CI for p.

Summary: Gaussian and CLT

> 1. **Gaussian:** $\mathcal{N}(\mu, \sigma^2) : f_X(x) = \ldots$ "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.