ESE 402/542 Recitation 2: Probability Review

Last Time...

- Covered fundamental definitions of a random variable.
- PMF, PDF, and CDF.
- Functions of a random variable: expectation and variance.
- **Cautionary tales:** CDF always exists and characterizes distribution, but PDF may not exist (non-differentiable CDFs). Expectation and variance may not exist, even if distribution visually looks well-behaved (Cauchy distribution).

Topics for today

Topics for today

- Another identifier of a distribution: Moment-Generating Function.

Topics for today

- Another identifier of a distribution: Moment-Generating Function.
- The tour de force of statistics: Central Limit Theorem.

Topics for today

- Another identifier of a distribution: Moment-Generating Function.
- The tour de force of statistics: Central Limit Theorem.
- The random variable zoo – basic definitions and properties.

Moment-Generating Function

- Mathematical definition: Given random variable X, MGF is

$$M_X(t) = \mathbb{E}\left[e^{tX}\right]$$

Moment-Generating Function

- Mathematical definition: Given random variable X, MGF is

$$M_X(t) = \mathbb{E}\left[e^{tX}\right]$$

- A form of ID: if two random variables X, Y have satisfy $M_X(t) = M_Y(t)$ for all t, then X and Y are identical.

Moment-Generating Function

- Mathematical definition: Given random variable X, MGF is

$$M_X(t) = \mathbb{E}\left[e^{tX}\right]$$

- A form of ID: if two random variables X, Y have satisfy $M_X(t) = M_Y(t)$ for all t, then X and Y are identical.
- Caveats: MGF may not exist for all t, or at all.

Moment-Generating Function

- Why is it called “moment-generating function”?

Moment-Generating Function

- Why is it called "moment-generating function"?
  - First observe by Taylor series and linearity of expectation that

$$\begin{aligned} \mathbb{E}\left[e^{tX}\right] &= \mathbb{E}\left[1 + tX + \frac{t^2}{2!}X^2 + \cdots\right] \\ &= 1 + t\mathbb{E}[X] + \frac{t^2}{2!}\mathbb{E}\left[X^2\right] + \cdots \end{aligned}$$

Moment-Generating Function

- Why is it called “moment-generating function”?
  - First observe by Taylor series and linearity of expectation that

$$\begin{aligned}\mathbb{E}\left[e^{tX}\right] &= \mathbb{E}\left[1 + tX + \frac{t^2}{2!}X^2 + \cdots\right] \\ &= 1 + t\mathbb{E}\left[X\right] + \frac{t^2}{2!}\mathbb{E}\left[X^2\right] + \cdots\end{aligned}$$

  - Given random variable X, its first, second, third etc. moments are defined as $\mathbb{E}\left[X\right]$, $\mathbb{E}\left[X^2\right]$, $\mathbb{E}\left[X^3\right]$.

Moment-Generating Function

- Why is it called "moment-generating function"?
  - First observe by Taylor series and linearity of expectation that

$$\begin{aligned}\mathbb{E}\left[e^{tX}\right] &= \mathbb{E}\left[1 + tX + \frac{t^2}{2!}X^2 + \cdots\right] \\ &= 1 + t\mathbb{E}\left[X\right] + \frac{t^2}{2!}\mathbb{E}\left[X^2\right] + \cdots\end{aligned}$$

  - Given random variable X, its first, second, third etc. moments are defined as $\mathbb{E}\left[X\right],\ \mathbb{E}\left[X^2\right],\ \mathbb{E}\left[X^3\right]$.
  - In math, "(exponential) generating function" of infinite sequence $\{a_n\}$ is

$$\sum_{n=0}^{\infty} a_n \frac{t^n}{n!}$$

Moment-Generating Function

- Why is it called "moment-generating function"?
  - First observe by Taylor series and linearity of expectation that

$$\begin{aligned}\mathbb{E}\left[e^{tX}\right] &= \mathbb{E}\left[1 + tX + \frac{t^2}{2!}X^2 + \cdots\right] \\ &= 1 + t\mathbb{E}\left[X\right] + \frac{t^2}{2!}\mathbb{E}\left[X^2\right] + \cdots\end{aligned}$$

  - Given random variable X, its first, second, third etc. moments are defined as $\mathbb{E}\left[X\right], \mathbb{E}\left[X^2\right], \mathbb{E}\left[X^3\right]$.
  - In math, "(exponential) generating function" of infinite sequence $\{a_n\}$ is

$$\sum_{n=0}^{\infty} a_n \frac{t^n}{n!}$$

  - Putting these together, a "moment" generating function is precisely

$$\sum_{n=0}^{\infty} \mathbb{E}\left[X^n\right] \frac{t^n}{n!} = \mathbb{E}\left[e^{tX}\right]$$

Discuss in Groups:

Given the MGF $M_X(t)$, how can you recover the k-th moment $\mathbb{E}\left[X^k\right]$?

Remember:

$$M_X(t) = 1 + t\mathbb{E}[X] + \frac{t^2}{2!}\mathbb{E}\left[X^2\right] + \cdots$$

Discuss in Groups:

Given the MGF $M_X(t)$, how can you recover the k-th moment $\mathbb{E}\left[X^k\right]$?

Remember:

$$M_X(t) = 1 + t\mathbb{E}[X] + \frac{t^2}{2!}\mathbb{E}\left[X^2\right] + \cdots$$

Solution:

$$\mathbb{E}\left[X^k\right] = \frac{d^k M_X(t)}{dt^k}\Big|_{t=0}$$

Taking the derivative removes the first $k-1$ terms, and evaluating at $t=0$ removes terms after k. The k-th term is exactly $\mathbb{E}\left[X^k\right]$.

Central Limit Theorem

- One of the most useful tools in statistical testing.

Central Limit Theorem

- One of the most useful tools in statistical testing.
- Long story short: the distribution of sample means converges to a normal distribution

Central Limit Theorem

- One of the most useful tools in statistical testing.
- Long story short: the distribution of sample means converges to a normal distribution
- VERY IMPORTANT CAVEAT: only applicable when variance is finite $\mathrm{var}(X) = \mathbb{E}\left[(X-\mu)^2\right] = \sigma^2 < \infty$

Central Limit Theorem

- One of the most useful tools in statistical testing.
- Long story short: the distribution of sample means converges to a normal distribution
- VERY IMPORTANT CAVEAT: only applicable when variance is finite $\text{var}(X) = \mathbb{E}\left[(X-\mu)^2\right] = \sigma^2 < \infty$
- Counterexample when variance doesn't exist: Cauchy distribution.

Central Limit Theorem

- Mathematically, given sample mean $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, where X_i are iid with mean μ and variance $\sigma^2 < \infty$, we have

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

$$\bar{X}_n - \mu \xrightarrow{d} \mathcal{N}(0, \frac{\sigma^2}{n})$$

Central Limit Theorem

- Mathematically, given sample mean $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, where X_i are iid with mean μ and variance $\sigma^2 < \infty$, we have

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$
$$\bar{X}_n - \mu \xrightarrow{d} \mathcal{N}(0, \frac{\sigma^2}{n})$$

- This essentially implies that when we increase the number of samples n used to compute the sample mean $\bar{X}_n$, the variance of the sample mean decreases inversely with n.

CLT: Statistical Testing

- How to use this? Given objects coming from an unknown or complicated distribution, but we have statistics like the mean μ and variance σ^2, we can use the normal distribution in place of the true distribution.

CLT: Statistical Testing

- How to use this? Given objects coming from an unknown or complicated distribution, but we have statistics like the mean μ and variance σ^2, we can use the normal distribution in place of the true distribution.
- Given a batch of objects that we want to decide comes from our population distribution, we can calculate its mean and compare with $\mathcal{N}(\mu, \frac{\sigma^2}{n})$ to see the probability it has that mean by chance–this is the beginning of statistical testing.

CLT: Statistical Testing

- How to use this? Given objects coming from an unknown or complicated distribution, but we have statistics like the mean μ and variance σ^2, we can use the normal distribution in place of the true distribution.
- Given a batch of objects that we want to decide comes from our population distribution, we can calculate its mean and compare with $\mathcal{N}(\mu, \frac{\sigma^2}{n})$ to see the probability it has that mean by chance–this is the beginning of statistical testing.
- If the batch has mean very far away from the population mean, then it is very very unlikely that it is like that by chance...we may deduce there might be something else other than random chance causing a systematic shift.

CLT: Statistical Testing

- How to use this? Given objects coming from an unknown or complicated distribution, but we have statistics like the mean μ and variance σ^2, we can use the normal distribution in place of the true distribution.
- Given a batch of objects that we want to decide comes from our population distribution, we can calculate its mean and compare with $\mathcal{N}(\mu, \frac{\sigma^2}{n})$ to see the probability it has that mean by chance–this is the beginning of statistical testing.
- If the batch has mean very far away from the population mean, then it is very very unlikely that it is like that by chance...we may deduce there might be something else other than random chance causing a systematic shift.
- Of course, the batch could come from a different distribution but have the same mean. How do we test this? Welcome to the zoo of statistical inference.

Discuss in Groups

Suppose we have 100 packages whose weights are *independent* and *uniformly* distributed between 5 and 50 lbs. What is the (approximate) probability that the total weight will exceed 3000 lbs? Give result in terms of the normal CDF.

Hint: If $U \sim \mathrm{Uniform}([a, b])$, $\mu = \frac{a+b}{2}$, $\sigma^2 = \frac{(b-a)^2}{12}$

Discuss in Groups

Suppose we have 100 packages whose weights are *independent* and *uniformly* distributed between 5 and 50 lbs. What is the (approximate) probability that the total weight will exceed 3000 lbs? Give result in terms of the normal CDF.

Hint: If $U \sim \mathrm{Uniform}([a, b])$, $\mu = \frac{a+b}{2}$, $\sigma^2 = \frac{(b-a)^2}{12}$

Solution: Let $S_{100} = \sum_{i=1}^{100} U_i$, where $U_i \sim \mathrm{Uniform}([5, 50])$.

$$
\begin{aligned}
P(S_{100} \leq 3000) &= P\left(\frac{S_{100} - 100 * \mu}{\sigma\sqrt{100}} \leq \frac{3000 - 100 * \mu}{\sigma\sqrt{100}}\right) \\
&\approx \Phi\left(\frac{3000 - 100 * \mu}{\sigma\sqrt{100}}\right) = \Phi(1.92)
\end{aligned}
$$

So $P(S_{100} > 3000) = 1 - \Phi(1.92)$