STAT 542: Summary to date
Where we have been
& where we are headed

- Completed
  - Intro to probability
  - Random variables (definition, cdf, pdf/pmf)
  - Transformations (introduction)
  - Expectation (moments, properties, mgf)
  - Inequalities (Markov, Chebychev, Jensen's)

- Next
  - Common univariate distributions
    * discrete examples (binomial, Poisson, geometric, hypergeomtric)
    * continuous examples (gamma, beta, normal)
    * exponential families, location-scale families
  - Multivariate Distributions

Common univariate distributions

Introduction

- Often it is useful to consider structural forms for a pdf f or cdf F, especially for modeling a population

- In particular, it is possible to specify a family of distributions using a single functional form with one or more free **parameters**.

- Examples
  - discrete:
  $$f_X(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$
  is a pmf for any $\lambda > 0$
  - continuous:
  $$f_X(x) = \frac{1}{\theta}e^{-x/\theta}, \quad x > 0$$
  is a pdf for any $\theta > 0$

- Parameters often correspond to characteristics of population being modeled
  e.g., both λ and θ correspond to the mean of each distribution above

Common univariate distributions

Discrete distributions: *discrete* uniform distribution

A Discrete Uniform $[1, N]$ random variable X has pmf

$$\mathbb{P}(X=x) = f(x) = \begin{cases} \frac{1}{N} & x = 1, 2, \ldots, N \\ 0 & \text{otherwise} \end{cases}$$

which models the random selection of an integer from 1 to N

To compute the mean/variance, one has resurrect formulas for sums of the first N integers (or sum of their squares)

$$\sum_{k=1}^{N} k = N(N+1)/2, \qquad \sum_{k=1}^{N} k^2 = N(N+1)(2N+1)/6. \quad (**)$$

$$\mathbb{E}(X) = \sum_{x=1}^{N} \frac{1}{N} x = \frac{1}{N} \sum_{x=1}^{N} x = \frac{1}{\cancel{N}} \frac{\cancel{N}(N+1)}{2} = \frac{N+1}{2} \quad \mathrm{I}$$

$$\mathbb{E}(X^2) = \sum_{x=1}^{N} \frac{x^2}{N} = \frac{1}{N} \sum_{x=1}^{N} x^2 \overset{(**)}{=} \frac{1}{\cancel{N}} \frac{\cancel{N}(N+1)(2N+1)}{6} = \frac{(N+1)(2N+1)}{6} \quad \mathrm{II}$$

$\mathbb{E}(X^2)$

$$\mathrm{Var}(X) \overset{\text{def}}{=} \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \overset{\mathrm{I}\,\&\,\mathrm{II}}{=} \frac{(N+1)(2N+1)}{6} - \left(\frac{N+1}{2}\right)^2 = \frac{(N+1)(N-1)}{12} \quad (*)$$

$X \sim \mathrm{Uni}(1, N)$

Parameters

If one wants Y to be discrete uniform$[N_0, N_1]$, just take $X \sim$discrete uniform$[1, N_1 - N_0 + 1]$ and define $\boxed{Y = X + N_0 - 1}$

$$\mathbb{E}(Y) = \mathbb{E}(X) + N_0 - 1 = \frac{\cancel{1} + N_1 - N_0 + \cancel{1} + 2N_0 - \cancel{2}}{2} = \frac{N_1 + N_0}{2}$$

$\mathbb{E}(Y)$ when $Y \sim \mathrm{Uni}[N_0, N_1]$

$$\mathrm{Var}(Y) = \mathrm{Var}(X + N_0 - 1) \underset{X \sim \mathrm{Uni}(1, N_1 - N_0 + 1)}{=} \frac{(N_1 - N_0 + 1 + 1)(N_1 - N_0 + \cancel{1} - \cancel{1})}{12} = \frac{(N_1 - N_0 + 2)(N_1 - N_0)}{12}$$

Common univariate distributions

Discrete distributions: Bernoulli distribution

Parameter

Bernoulli trial $X \sim$Bern(p)

- Two outcomes usually labeled 1, 0 (success/failure)

$$X = \begin{cases} 0 & \text{with prob } 1-p \\ 1 & \text{with Prob. } p \end{cases}$$

- p is a parameter $0 < p < 1$ (i.e., X is 1 with probability p)

If $X \sim Ber(p)$, $\mathbb{P}(X=0)=1-p$, $\mathbb{P}(X=1)=p$

- $f_X(0) = 1 - p$ and $f_X(1) = p$, or equivalently

$$f_X(x) = f_X(x|p) = P(X = x|p) = p^x(1-p)^{1-x}, \quad x = 0, 1$$

If $X \sim Ber(p)$, $\mathbb{P}(X=x)=p^x(1-p)^{1-x}$

$$\sum_{x=0}^{1} x\,\mathbb{P}(X=x)$$

✓ - $\text{E}X = \underset{x}{(0)}\underbrace{(1-p)}_{p(0)} + \underset{x}{(1)}\underset{p(1)}{(p)} = p$

$$\mathbb{E}(X^2) = \sum_{x=0}^{1} x^2\,\mathbb{P}(X=x) = \cancel{(0)^2(1-p)} + 1\,p = p \rightarrow \mathbb{E}X^2$$

- $\text{Var}(X) = \text{E}X^2 - (\text{E}X)^2 = p - p^2 = p(1-p)$

$$\text{If } X \sim Ber(p) \Rightarrow M_X(t) \stackrel{\text{def}}{=} \mathbb{E}[e^{tX}] = \sum_{x=0}^{1} e^{tx}\,\mathbb{P}(X=x)$$

$$= \cancel{e^0}(1-p) + e^t p = \boxed{1-p+e^t p}$$

MGF

Common univariate distributions

Discrete distributions: Binomial distribution

Binomial distribution, $X \sim$Binom(n, p), $0 < p < 1$

trails

- pmf given by

$$\mathbb{P}(X=x) = f_X(x) = f_X(x|n,p) = P(X = x|n,p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n$$

- Motivation: distribution for the number of successes in n independent Bernoulli(p) trials, i.e., if $Y_1, \ldots, Y_n$ are $\boxed{\text{independent}}$ Bern(p), where Y_i is the outcome of the ith trial ($Y_i = 1$ if the trial is "success" and 0 if "failure"), then $X = \sum_{i=1}^n Y_i$ is Binom(n, p) distributed

you need to show

def of X $\quad \mathbb{P}(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$

For a given $x = 0, 1, \ldots, n$, $P(X = x) = P\left(\sum_{i=1}^n Y_i = x\right)$

$Y_1 \ldots Y_n$ are independent $\qquad$ Bern(p)

$$= \sum_{\substack{(y_1,\ldots,y_n)\in\{0,1\}^n,\\ \sum_{i=1}^n y_i = x}} P(Y_1 = y_1, \ldots, Y_n = y_n) = \sum_{\substack{(y_1,\ldots,y_n)\in\{0,1\}^n,\\ \sum_{i=1}^n y_i = x}} \prod_{i=1}^n P(Y_i = y_i)$$

$$= \sum_{\substack{(y_1,\ldots,y_n)\in\{0,1\}^n,\\ \sum_{i=1}^n y_i = x}} \underbrace{\prod_{i=1}^n p^{y_i}(1-p)^{1-y_i}}_{p^{\sum y_i}(1-p)^{n-\sum y_i}} = \sum_{\substack{(y_1,\ldots,y_n)\in\{0,1\}^n,\\ \sum_{i=1}^n y_i = x}} p^x (1-p)^{n-x}$$

$$= p^x (1-p)^{n-x} \times \text{“}\underbrace{\textit{\# of ways to choose exactly } x \textit{ components of } (y_1, \ldots, y_n) \textit{ to be 1}}_{\binom{n}{x}}\text{”}$$

- mean: $\mathrm{E}X = np$ (proof next slide)

- variance: $\mathrm{Var}(X) = np(1-p)$

- Moment generating function: $M_X(t) = \mathrm{E}e^{tX} = [pe^t + (1-p)]^n$ for any $t \in \mathbb{R}$

  We've actually proven this already in the mgf section; the result follows from $M_X(t) = \sum_{x=0}^n \binom{n}{x}(pe^t)^x(1-p)^{n-x}$ and the binomial formula for $(a+b)^n$

Common univariate distributions

$X=\sum_{i=1}^{n} Y_i$, $Y_i \sim Ber(p)$, Y_i are ind.

Discrete distributions: Binomial distribution (cont'd)

Proof: $M_X(t) \overset{def}{=} \mathbb{E}\left[e^{tX}\right] = \mathbb{E}\left[e^{t(Y_1+Y_2+\cdots+Y_n)}\right] = \mathbb{E}\left[e^{tY_1}\cdot e^{tY_2}\cdots e^{tY_n}\right] = \mathbb{E}\left[e^{tY_1}\right]\cdots\mathbb{E}\left[e^{tY_n}\right] = M_{Y_1}(t)\cdots M_{Y_n}(t) = (1-p+pe^t)\cdots(1-p+pe^t) = \boxed{(1-p+pe^t)^n}$

Derive mean

1. Using mgf $M_X(t) = \mathrm{E}e^{tX} \overset{?}{=} [pe^t + (1-p)]^n$,

$$\left.\frac{d[pe^t + (1-p)]^n}{dt}\right|_{t=0} = n[pe^t + (1-p)]^{n-1}pe^t\Big|_{t=0} = np$$

2. Or let $X = \sum_{i=1}^n Y_i$ where $Y_1, \ldots, Y_n$ are independent Bern(p), so that

$$\mathrm{E}X = \mathrm{E}\left(\sum_{i=1}^n Y_i\right) = \sum_{i=1}^n \mathrm{E}Y_i = \sum_{i=1}^n p = np$$

3. Or using the direct definition

$$\begin{aligned}
\mathrm{E}X \overset{def}{=} \sum_{x=0}^n \underbrace{x\binom{n}{x}p^x(1-p)^{n-x}}_{\mathbb{P}(X=x)} &= \sum_{x=1}^n x\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} \\
&= \sum_{x=1}^n x\frac{n}{x}\frac{(n-1)!}{(x-1)!(n-x)!}p^x(1-p)^{n-x} \\
&= n\sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!}p^x(1-p)^{n-x} \\
&= n\sum_{x=1}^n \binom{n-1}{x-1}p^x(1-p)^{n-x} \\
&\overset{z=x-1}{=} n\sum_{z=0}^{n-1}\binom{n-1}{z}p^{z+1}(1-p)^{n-(z+1)} \qquad (z = x-1) \\
&= np\underbrace{\sum_{z=0}^{n-1}\binom{n-1}{z}p^z(1-p)^{n-1-z}}_{=1}
\end{aligned}$$

pmf of a Bin $(n-1, p)$

Variance derivation is similar but messy: compute

$$\mathrm{E}X(X-1) = \sum_{x=0}^n x(x-1)\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=2}^n \frac{n!}{(x-2)!(n-2)!}p^x(1-p)^{n-x} = n(n-1)p^2$$

Then, $\mathrm{Var}(X) = \mathrm{E}X(X-1) + \mathrm{E}X - (\mathrm{E}X)^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)$

$\mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}(X))^2$

Note: $\left.\frac{d^2M_X(t)}{dt^2}\right|_{t=0} = \mathbb{E}X^2$