Lecture 3:
Normal Distribution (Continued);
Two useful Discrete Distributions:
Binomial and Poisson
Chapter 1
Back to Standard Normal (Z): backwards?

- If I give you a probability, can you find the corresponding z value?
  → called percentiles

  – What is the z-score for the 25\textsuperscript{th} percentile of the N(0,1) curve?
    -0.67
  – 90\textsuperscript{th} percentile?
    1.68
Standardizing

- We can convert any normal to a standard normal distribution
- To do this, just subtract the mean and divide by the standard deviation
- z-score – standardized value of x (how many standard deviations from the mean)

\[ z = \frac{x - \mu}{\sigma} \]
Standardizing

• Put differently...
• Suppose we want the area between $a$ and $b$ for $x$
• This is exactly the same area between $a^*$ and $b^*$ for $z$, where $a^*$ is the $a$ standardized and $b^*$ is $b$ standardized

$$\int_{a}^{b} f(x) \, dx = \int_{a^*}^{b^*} f(z) \, dz$$
Standard Normal Distribution

• The **standardized values** for any distribution always have mean 0 and standard deviation 1.
• If the original distribution is normal, the standardized values have normal distribution with mean 0 and standard deviation 1.
• Hence, the **standard normal distribution** is extremely important, especially it’s corresponding Z table.
  – Remember we can do this forward or backward (using percentiles)
Practice

• In 2000 the scores of students taking SATs were approximately normal with mean 1019 and standard deviation 209. What percent of all students had the SAT scores of:
  – at least 820? (limit for Division I athletes to compete in their first college year)
    82.89%
  – between 720 and 820? (partial qualifiers)
    9.47%
  – How high must a student score in order to place in the top 20% of all students taking the SAT?
    1195
  – Berry’s score was the 68th percentile, what score did he receive?
    1117
Some Useful Discrete Distributions

• Binomial Distribution

• Poisson Distribution

• Poisson Approximation to the Binomial Distribution
1.6
Binomial Distribution

• Tossing an unfair coin (many times). Suppose the proportion of times we get Heads (chances of getting H) is \( \pi \), and the proportion of times we get Tails (chances of getting T) is \((1-\pi)\), \(0<\pi<1\).

• Now if we look at 3 independent tosses.
  – Independent
    • chances that the second toss is H does not depend on the first toss, etc
Binomial probabilities

• The probabilities associated with each outcome are as follows:

\[
P(\text{HHH}) = \pi \pi \pi = \pi^3
\]

\[
P(\text{HHT}) = \pi \pi (1-\pi) = \pi^2 (1-\pi)
\]

\[
P(\text{HTH}) = \pi (1-\pi) \pi = \pi^2 (1-\pi)
\]

\[
P(\text{HTT}) = \pi (1-\pi)(1-\pi) = (1-\pi)^2 \pi
\]

\[
P(\text{THH}) = (1-\pi) \pi \pi = \pi^2 (1-\pi)
\]

\[
P(\text{THT}) = (1-\pi) \pi (1-\pi) = (1-\pi)^2 \pi
\]

\[
P(\text{TTH}) = (1-\pi)(1-\pi) \pi = (1-\pi)^2 \pi
\]

\[
P(\text{TTT}) = (1-\pi)(1-\pi)(1-\pi) = (1-\pi)^3
\]

• Let \( X \) denote the number of H in these 3 tosses, then \( X \) is a discrete variable. Let’s look at the mass function...
Binomial cont.

• Mass function of $X$ (# of H in 3 tosses):
  
  $p(X=0) = (1-\pi)^3 = \frac{3!}{0!(3-0)!}(1-\pi)^3$

  $p(X=1) = 3(1-\pi)^2\pi = \frac{3!}{1!(3-1)!}(1-\pi)^2\pi$

  $p(X=2) = 3\pi^2(1-\pi) = \frac{3!}{2!(3-2)!}\pi^2(1-\pi)$

  $p(X=3) = \pi^3 = \frac{3!}{3!(3-3)!}\pi^3$

• In general, we can write the mass function

  $p(X = x) = p(x) = \frac{3!}{x!(3-x)!} \pi^x (1-\pi)^{3-x}$, \hspace{1cm} x = 0, 1, 2, 3
More generally, if we toss the coin \( n \) times and let \( X \) be the number of heads in \( n \) tosses, the mass function of \( X \) would be:

\[
p(X = x) = p(x) = \frac{n!}{x!(n-x)!} \pi^x (1-\pi)^{n-x}, \quad x = 0, 1, \ldots, n
\]

Here \( X \) is said to have a binomial distribution. Used to count number of “successes” in \( n \) “trials.” Table II gives \( p(x) \) for some selected \( n \) and \( \pi \).
Example

• Suppose we know that for a particular brand of light bulb, the chance that a randomly selected light bulb is defected is 0.1 (i.e., defect rate is 0.1). Now we check a batch of 5 of such light bulbs
  – What is the chance that we get no defected ones?
  – What is the chance that we get less than 3 defected ones?
Use Table II for Binomial Distributions

- $n = 5$ and $\pi = 0.1$, so we can use Table II. Here's what it looks like.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$\pi$</th>
<th>$\pi$</th>
<th>$\pi$</th>
<th>$\pi$</th>
<th>$\pi$</th>
<th>$\pi$</th>
<th>$\pi$</th>
<th>$\pi$</th>
<th>$\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.774</td>
<td>.590</td>
<td>.328</td>
<td>.237</td>
<td>.168</td>
<td>.078</td>
<td>.031</td>
<td>.010</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>.203</td>
<td>.329</td>
<td>.409</td>
<td>.396</td>
<td>.360</td>
<td>.259</td>
<td>.157</td>
<td>.077</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>.022</td>
<td>.072</td>
<td>.205</td>
<td>.263</td>
<td>.309</td>
<td>.346</td>
<td>.312</td>
<td>.230</td>
<td>...</td>
</tr>
<tr>
<td>3</td>
<td>.001</td>
<td>.009</td>
<td>.051</td>
<td>.088</td>
<td>.132</td>
<td>.230</td>
<td>.312</td>
<td>.346</td>
<td>...</td>
</tr>
<tr>
<td>4</td>
<td>.000</td>
<td>.000</td>
<td>.007</td>
<td>.015</td>
<td>.029</td>
<td>.077</td>
<td>.157</td>
<td>.259</td>
<td>...</td>
</tr>
<tr>
<td>5</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.001</td>
<td>.002</td>
<td>.010</td>
<td>.031</td>
<td>.078</td>
<td>...</td>
</tr>
</tbody>
</table>
Q1: the chance of no defected bulbs...

- No defected ones, so \( x = 0 \).

\[
\begin{array}{cccccccccc}
  \pi \\
x & .05 & .1 & .2 & .25 & .3 & .4 & .5 & .6 & \ldots \\
0 & .774 & .590 & .328 & .237 & .168 & .078 & .031 & .010 & \ldots \\
1 & .203 & .329 & .409 & .396 & .360 & .259 & .157 & .077 & \ldots \\
2 & .022 & .072 & .205 & .263 & .309 & .346 & .312 & .230 & \ldots \\
3 & .001 & .009 & .051 & .088 & .132 & .230 & .312 & .346 & \ldots \\
4 & .000 & .000 & .007 & .015 & .029 & .077 & .157 & .259 & \ldots \\
5 & .000 & .000 & .000 & .001 & .002 & .010 & .031 & .078 & \ldots \\
\end{array}
\]

Therefore (proportion with \( x = 0 \)) = \( p(0) = .590 \)
Q2: less than 3 defected bulbs

- Less than 3: $x = 0, 1, 2$.

| $x$ | .05 | .1  | .2  | .25 | .3  | .4  | .5  | .6  |...
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0   | .774| .590| .328| .237| .168| .078| .031| .010|...
| 1   | .203| .329| .409| .396| .360| .259| .157| .077|...
| 2   | .022| .072| .205| .263| .309| .346| .312| .230|...
| 3   | .001| .009| .051| .088| .132| .230| .312| .346|...
| 4   | .000| .000| .007| .015| .029| .077| .157| .259|...
| 5   | .000| .000| .000| .001| .002| .010| .031| .078|...

(proportion with $x < 3) = p(0) + p(1) + p(2) = .590 + .329 + .072 = .991
Poisson Distribution

• Used to model the # of times an event occurs during a particular time period (or region of space), such as:
  – $X$: # of car accidents that happen in 30 days
  – $X$: # of calls a receptionist receives in an hour, etc
• The Poisson mass function is:

$$p(X = x) = p(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \ldots.$$ 

where $\lambda$ must be positive.

• Appendix Table III gives a small tabulation of Poisson mass function for selected values of $\lambda$. 
Average rate for Poisson

- Here, \( \lambda \) can be depicted as the “average rate” of arrival.

- For example, let’s say a receptionist receives an average of 5 calls every hour, then if \( X \) denotes the # of calls the receptionist gets during the next hour, \( X \) follows a Poisson distribution with \( \lambda = 5 \).

- Using this example...
  - What is the chance the receptionist receives more than 2 calls during the next hour?

\[
(\text{proportion with } x > 2) = 1 - (\text{proportion with } x \leq 2) \\
= 1 - [p(0) + p(1) + p(2)] \\
= 1 - .125 = .875
\]
Review

• How to graphically display data
  – Histograms, dot plots, stem plots, etc
  – Helps to show how samples are distributed

• Distributions of both continuous and discrete variables
  – Density functions and Mass functions
  – Shows the distribution of the entire population or process

• Some important distributions
  – Continuous: Exponential, Normal, Uniform ...
  – Discrete: Binomial, Poisson ...
Density (for Continuous) and Mass (for Discrete) functions

• tell you the “chance/proportion/probability” that a variable takes a certain value
  – Need to know the distribution expression

• both used to rigorously describe populations or processes
  – How to know which distribution is applicable?

→ See Chapter 2
  • Numerical measures for both samples and populations
  • Bring Your Calculator from now on...
When You Go Home...

- Get The Textbook
  - *bring a calculator and the tables to every class*

- Review Ch. 1 (1.1 – 1.4 and 1.6), start Hw#2

- Read Sections 2.1 (till trimmed means, Pg 63) and 2.2 (till sample variances, top of Pg 74)