UCLA STAT 110 A

Applied Probability & Statistics for Engineers

●**Instructor:** Ivo Dinov,

Asst. Prof. In Statistics and Neurology

●**Teaching Assistant:** Maria Chang, UCLA Statistics

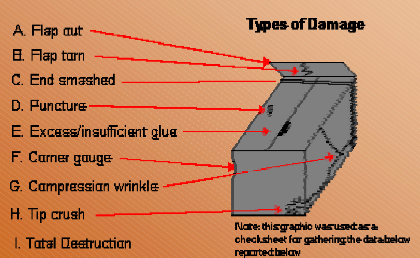
University of California, Los Angeles, Spring 2003

http://www.stat.ucla.edu/~dinov/

Chapters 3 – Discrete Variables, Probabilities, CLT

- Random Variables (RV's
- Probability Density Functions (PDF's) for discrete RV's
- Binomial, NegativeBinomial, Geometric,
- Hypergeometric, Poisson distributions

Frequency Distributions- damaged boxes



Frequency Distributions- damaged boxes

Type	Total Frequency	Relative Frequency	Percentage
A - Flap out	16	0.0096	1
B - Flap torn	17	0.0102	1
C - End smashed	132	0.0793	8
D – Puncture	95	0.0571	6
E - Glue problem	87	0.0523	5
F - Corner gouge	984	0.5913	59
G – Compr. wrinkle	15	0.0090	1
H - Tip crushed	303	0.1821	18
I - Tot. destruction	15	0.0090	1
Total	1664	0.9999*	100

(* the relative frequencies do not add to 1.0000 due to rounding)

Frequency Distributions- damaged boxes

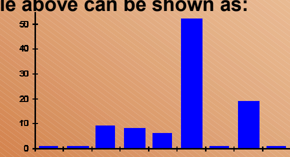
Relative frequency for type A is: $\frac{16}{1664} = 0.0096$

Percentage for type A is: $\frac{16}{1664} \times 100 = 0.96 \approx 1$ **percent.**

The usefulness of relative frequencies and percentages is clear: for example, it is easily seen that corner gouge accounts for 59% of the total number of damages.

Frequency Distributions- damaged boxes

The frequency distribution of a variable is often presented graphically as a bar-chart/bar-plot. For example, the data in the frequency table above can be shown as:



The vertical axis can be frequencies or relative frequencies or percentages. On the horizontal axis all boxes should have the same width leave gaps between the boxes (because there is no connection between them) the boxes can be in any order.

Experiments, Models, RV's

- An **experiment** is a naturally occurring phenomenon, a scientific study, a sampling trial or a test., in which an object (unit/subject) is selected at random (and/or treated at random) to *observe/measure* different outcome characteristics of the process the experiment studies.
- **Model** – generalized hypothetical description used to analyze or describe a phenomenon.
- A **random variable** is a type of measurement taken on the outcome of a random experiment.

Definitions

- The ***probability function*** for a discrete random variable ***X*** gives the chance that the observed value for the process equals a specific outcome, ***x***.
  - $P(X = x)$ [denoted $pr(x)$ or $P(x)$]

    for every value x that the R.V. X can take
- E.g., number of heads when a coin is tossed twice

x	0	1	2
pr(x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Stopping at one of each or 3 children

Sample Space – **complete/unique description of the possible outcomes from this experiment.**

Outcome	GGG	GGB	GB	BG	BBG	BBB
Probability	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

- For R.V. X = number of girls, we have

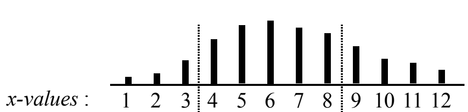
X	0	1	2	3
pr(x)	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Tossing a biased coin twice

- For each toss, P(Head) = p → P(Tail) = P(comp(H))=1-p
- Outcomes: HH, HT, TH, TT
- Probabilities: $p.p$, $p(1\text{-}p)$, $(1\text{-}p)p$, $(1\text{-}p)(1\text{-}p)$
- Count X, the number of heads in 2 tosses

X	0	1	2
pr(x)	$(1-p)^2$	$2p\ (1-p)$	p^2

Calculating Interval probabilities from cumulative probabilities



To get 4 to 8 — *P(3< X <9)*

start with everything up to 8 — *P(X <9)*

and remove from 3 down — *P(X<=3)*

How to find the upper-tail?

Bernoulli Trials

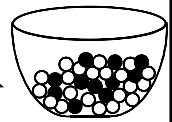
- A Bernoulli trial is an experiment where only two possible outcomes are possible (0 / 1).
- Examples:
  - Coin tosses
  - Computer chip (0 / 1) signal.
  - Poll supporters/opponents; yes/no; for/against.

The two-color urn model

N balls in an urn, of which there are

M black balls

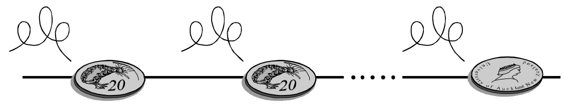
$N - M$ white balls



Sample n balls and count X = # black balls in sample

We will compute the probability distribution of the R.V. X

The biased-coin tossing model



toss 1 pr(H) = p — toss 2 pr(H) = p — ..... — toss n pr(H) = p

Perform n tosses and count X = # heads

We also want to compute the probability distribution of this R.V. X!
Are the two-color urn and the biased-coin models <u>related</u>? How do we present the models in <u>mathematical terms</u>?

The answer is: Binomial distribution

- The distribution of the number of heads in n tosses of a biased coin is called the ***Binomial distribution***.

Binomial(N, p) – the probability distribution of the number of Heads in an N-toss coin experiment, where the probability for Head occurring in each trial is p.
E.g., Binomial(6, 0.7)

	x	0	1	2	3	4	5	6
Individual	pr($X = x$)	0.001	0.010	0.060	0.185	0.324	0.303	0.118
Cumulative	pr($X \le x$)	0.001	0.011	0.070	0.256	0.580	0.882	1.000

For example P(X=0) = P(all 6 tosses are Tails) =

$$(1-0.7)^6 = 0.3^6 = 0.001$$

Binary random process

The ***biased-coin tossing model*** is a physical model for situations which can be characterized as a series of trials where:

- each trial has only two outcomes: *success* or *failure*;
- p = P(*success*) is the same for every trial; and
- trials are independent.

- The distribution of X = number of successes (heads) in N such trials is

Binomial(N, p)

Sampling from a finite population – Binomial Approximation

If we take a sample of size n

- from a much larger population (of size N)
- in which a proportion p have a characteristic of interest, then the distribution of ***X*, the number in the sample with that characteristic**,
- is *approximately* Binomial(n, p).
  - (Operating Rule: Approximation is adequate if $n / N < 0.1$.)
- Example, polling the US population to see what proportion is/has-been married.

Binomial Probabilities – the moment we all have been waiting for!

- Suppose X ~ Binomial(n, p), then the probability

$$P(X=x)=\binom{n}{x}p^{x}(1-p)^{(n-x)},\quad 0\le x\le n$$

- Where the binomial coefficients are defined by

$$\binom{n}{x}=\frac{n!}{(n-x)!\;x!},\qquad n!=1\times 2\times 3\times\ldots\times(n-1)\times n$$

n-factorial

Binomial Formula with examples

- Does the Binomial probability satisfy the requirements?

$$\Sigma_x P(X=x)=\Sigma_x\binom{n}{x}p^{x}(1-p)^{(n-x)}=(\mathrm{p}+(1-\mathrm{p}))^{\mathrm{n}}=1$$

- Explicit examples for n=2, do the case n=3 at home!

$$\sum_{x=0}^{2}\binom{2}{x}p^{x}(1-p)^{(2-x)}=$$ { *Three terms in the sum*

$$\binom{2}{0}p^{0}(1-p)^{2}+\binom{2}{1}p^{1}(1-p)^{1}+\binom{2}{2}p^{2}(1-p)^{0}=$$

$$1\times1\times(1-p)^{2}+2\times p\times(1-p)+1\times p^{2}\times1=$$

$$(p+(1\text{-}p))^{2}=1$$

{ *Usual quadratic-expansion formula*

Examples – Birthday Paradox

- The Birthday Paradox: In a random group of N people, what is the change that at least two people have the same birthday?
- E.x., if N=23, P>0.5. Main confusion arises from the fact that in real life we rarely meet people having the same birthday as us, and we meet more than 23 people.
- The reason for such high probability is that any of the 23 people can compare their birthday with any other one, not just you comparing your birthday to anybody else's.
- There are N-Choose-2 = 20*19/2 ways to select a pair or people. Assume there are 365 days in a year, P(one-particular-pair-same-B-day)=1/365, and
- P(one-particular-pair-failure)=1-1/365 ~ 0.99726.
- For N=20, 20-Choose-2 = 190. E={No 2 people have the same birthday is the event all 190 pairs fail (have different birthdays)}, then P(E) = P(failure)^{190} = 0.99726^{190} = 0.59.
- Hence, P(at-least-one-success)=1-0.59=0.41, quite high.
- Note: for N=42 ➔ P>0.9 …

Expected values

- The game of chance: cost to play:\$1.50; Prices {\$1, \$2, \$3}, probabilities of winning each price are {0.6, 0.3, 0.1}, respectively.
- Should we play the game? What are our chances of winning/loosing?

Prize (\$)	x	1	2	3	
Probability	pr(x)	0.6	0.3	0.1	
What we would "expect" from 100 games					*add across row*
Number of games won		0.6×100	0.3×100	0.1×100	↙
\$ won		$1\times0.6\times100$	$2\times0.3\times100$	$3\times0.1\times100$	Sum

Total prize money = Sum; ***Average prize money*** = Sum/100 $= 1\times0.6 + 2\times0.3 + 3\times0.1 = 1.5$

<u>Theoretically</u> Fair Game: price to play EQ the expected return!

Average Winnings from a Game conducted *N* times

Number of games played (N)	Prize won in dollars(x): 1	2	3	Average winnings per game ($\bar{x}$)
	frequencies (Relative frequencies)			
100	64 (.64)	25 (.25)	11 (.11)	1.7
1,000	573 (.573)	316 (.316)	111 (.111)	1.538
10,000	5995 (.5995)	3015 (.3015)	990 (.099)	1.4995
20,000	11917 (.5959)	6080 (.3040)	2000 (.1001)	1.5042
30,000	17946 (.5982)	9049 (.3016)	3005 (.1002)	1.5020
∞	(.6)	(.3)	(.1)	1.5

So far we looked at the theoretical expectation of the game. Now we simulate the game on a computer to obtain random samples from our distribution, according to the probabilities {0.6, 0.3, 0.1}.

Definition of the expected value, in general.

- The expected value:

$$E(X)=\sum_{\text{all } x} x\,P(x)\;\left(=\int_{\text{all } X} x\,P(x)dx\right)$$

- = Sum of (value times probability of value)

Example

In the at least one of each or at most 3 children example, where X ={number of Girls} we have:

X	0	1	2	3
pr(x)	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

$$E(X) = \sum_x x P(x)$$

$$= 0 \times \frac{1}{8} + 1 \times \frac{5}{8} + 2 \times \frac{1}{8} + 3 \times \frac{1}{8}$$

$$= 1.25$$

The expected value and population mean

μ_x = E(X) is called the ***mean*** of the distribution of X.

μ_X = E(X) is usually called the ***population mean.***

μ_x is the point where the bar graph of P($X = x$) balances.

Population standard deviation

The ***population standard deviation*** is

$$\text{sd}(X) = \sqrt{E[(X - \mu)^2]}$$

Note that if X is a RV, then (X-μ) is also a RV, and so is $(X-\mu)^2$. Hence, the *expectation*, $E[(X-\mu)^2]$, makes sense.

For the Binomial distribution . . . mean

E(X) = $n\,p$,

$$\text{sd}(X) = \sqrt{np(1-p)}$$

X~Binomial(n, p) ➔

$$X = Y_1 + Y_2 + Y_3 + .. + Y_n,$$

where $Y_k \sim Bernoulli(p)$,

$E(Y_1) = p$ ➔

$$E(X) = E(Y_1 + Y_2 + Y_3 + .. + Y_n) = np$$

Binomial and Multinomial Distributions

- **Multinomial Distribution**
  - k possible outcomes ($E_1, \cdots, E_k$)
  - Each outcome has probability p_i ($p_1 + \cdots + p_k = 1$)
  - In n **independent** trials, $X_1 + X_2 + \cdots + X_k = n$

$$f(x_1, \cdots, x_k; p_1, \cdots, p_k, n) = \binom{n}{x_1, x_2, \cdots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

$$\text{with } \sum_{i=1}^{k} x_i = n, \quad \sum_{i=1}^{k} p_i = 1$$

Marginal distribution of X_i: $Bin(n, p_i)$

Binomial and Multinomial Distributions

Ex. Suppose we have 9 people arriving at a meeting.

P(by Air) = 0.4, P(by Bus) = 0.2

P(by Automobile) = 0.3, P(by Train) = 0.1

P(3 by Air, 3 by Bus, 1 by Auto, 2 by Train) = ?

P(2 by air) = ?

Linear Scaling (affine transformations) $aX + b$

For any constants a and b, the expectation of the RV $aX + b$ is equal to the sum of the product of a and the expectation of the RV X and the constant b.

$$E(aX + b) = a\,E(X) + b$$

And similarly for the standard deviation (b, an additive factor, does not affect the SD).

$$SD(aX + b) = |a|\,SD(X)$$

Linear Scaling (affine transformations) $aX + b$

Why is that so?

$E(aX + b) = a\,E(X) + b \qquad SD(aX + b) = |a|\,SD(X)$

$$E(aX+b) = \sum_{x=0}^{n} (a x + b)\ P(X = x) =$$

$$\sum_{x=0}^{n} a\,x\,P(X=x) + \sum_{x=0}^{n} b\ P(X=x) =$$

$$a\sum_{x=0}^{n} x\,P(X=x) + b\sum_{x=0}^{n} P(X=x) =$$

$$aE(X) + b\times 1 = aE(X) + b.$$

Linear Scaling (affine transformations) $aX + b$

Example:

$E(aX + b) = a\,E(X) + b \qquad SD(aX + b) = |a|\,SD(X)$

1. X={-1, 2, 0, 3, 4, 0, -2, 1}; P(X=x)=1/8, for each x
2. Y = 2X-5 = {-7, -1, -5, 1, 3, -5, -9, -3}
3. E(X)=
4. E(Y)=
5. Does **E(X) = 2 E(X) –5 ?**
6. Compute SD(X), SD(Y). **Does SD(Y) = 2 SD(X)?**

Linear Scaling (affine transformations) $aX + b$

And why do we care?

$E(aX + b) = a\,E(X) + b \qquad SD(aX + b) = |a|\,SD(X)$

-completely general strategy for computing the distributions of RV's which are obtained from other RV's with known distribution. E.g., X~N(0,1), and Y=aX+b, then we need **not** calculate the mean and the SD of Y. We know from the above formulas that E(Y) = b and SD(Y) =|a|.

-These formulas hold for all distributions, not only for Binomial and Normal.

Linear Scaling (affine transformations) $aX + b$

And why do we care?

$E(aX + b) = a\,E(X) + b \qquad SD(aX + b) = |a|\,SD(X)$

-E.g., say the rules for the game of chance we saw before change and the new pay-off is as follows: {\$0, \$1.50, \$3}, with probabilities of {0.6, 0.3, 0.1}, as before. What is the newly expected return of the game? Remember the old expectation was equal to the entrance fee of \$1.50, and the game was fair!

Y = 3(X-1)/2

{\$1, \$2, \$3} → {\$0, \$1.50, \$3},

E(Y) = 3/2 E(X) –3/2 = 3 / 4 = \$0.75

And the game became clearly biased. Note how easy it is to compute E(Y).

Means and Variances for (in)dependent Variables!

- **Means**:
  - Independent/Dependent Variables {X1, X2, X3, …, X10}
    - E(X1 + X2 + X3 + … + X10) = E(X1)+ E(X2)+ E(X3)+… + E(X10)
- **Variances**:
  - Independent Variables {X1, X2, X3, …, X10}, variances add-up
    Var(X1 +X2 + X3 + … + X10) = Var(X1)+Var(X2)+Var(X3)+…+Var(X1)
  - Dependent Variables {X1, X2}
    Variance contingent on the variable dependences,
    - E.g., If X2 = 2X1 + 5,

Var(X1 +X2) =Var (X1 + 2X1 +5) =

Var(3X1 +5) =Var(3X1) = 9Var(X1)

For the Binomial distribution . . . SD

E(***X***) = ***n p***

$SD(X) = \sqrt{np(1-p)}$

X~Binomial(n, p) ➔
$X = Y_1 + Y_2 + Y_3 + \ldots + Y_n$,
where $Y_k \sim$ Bernoulli(p),
$Var(Y_1) = (1-p)^2 \times p + (0-p)^2 \times (1-p)$ ➔
$Var(Y_1) = (1-p)(p-p^2+p^2) = (1-p)p$ ➔
$Var(X) = Var(Y_1) + \ldots + Var(Y_n) = n(1-p)p$
SD(X)=Sqrt[Var(X)] = Sqrt[n(1-p)p]

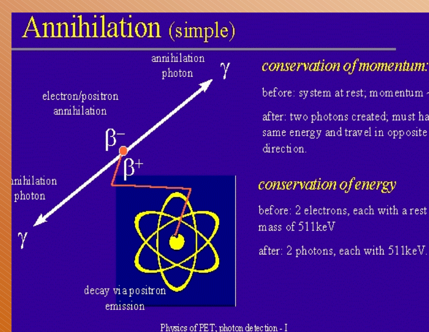
Sample vs. theoretical mean & varaince

- The **Expected value**: (population mean) $E(X) = \sum_{\text{all } x} x\, P(x) \quad \left(= \int_{\text{all } x} x\, P(x)dx \right)$
- **Sample mean** $\overline{X} = \frac{1}{N}\sum_{k=1}^{N} x_k$
- **(Theoretical) Variance**
  $$Var(X) = \sum_{\text{all } x} (x - \mu_x)^2 P(x) \quad \left(= \int_{\text{all } x} (x-\mu_x)^2 P(x)dx \right)$$
- **(Sample) variance**
  $$Var(X) = \frac{1}{N-1}\sum_{k=1}^{N}(x_k - \overline{X})^2 = \sum_{k=1}^{N}(x_k - \overline{X})^2 P(x)$$

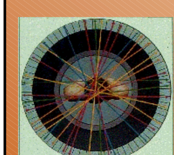
Poisson Distribution – Definition

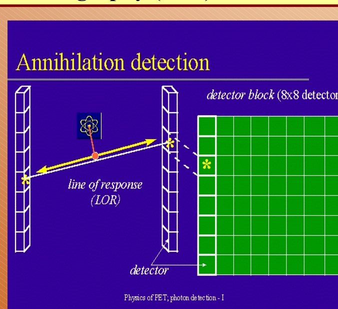
- Used to model counts – number of arrivals (k) on a given interval …
- The Poisson distribution is also sometimes referred to as the **distribution of rare events**. Examples of Poisson distributed variables are number of accidents per person, number of sweepstakes won per person, or the number of catastrophic defects found in a production process.

Functional Brain Imaging – Positron Emission Tomography (PET)



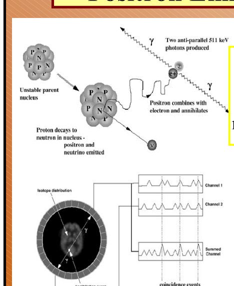
Functional Brain Imaging - Positron Emission Tomography (PET)





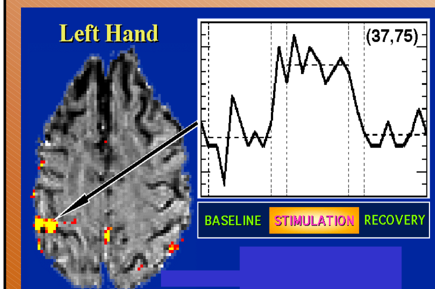
http://www.nucmed.buffalo.edu

Functional Brain Imaging – Positron Emission Tomography (PET)



Isotope	Energy (MeV)	Range(mm)	1/2-life	Appl.
^{11}C	0.96	1.1	20 min	receptors
^{15}O	1.7	1.5	2 min	stroke/activation
^{18}F	0.6	1.0	110 min	neurology
^{124}I	~2.0	1.6	4.5 days	oncology

Functional Brain Imaging – Positron Emission Tomography (PET)



Poisson Distribution – Mean

- Used to model counts – number of arrivals (k) on a given interval …
- Y~Poisson(λ), then P(Y=k) = $\frac{\lambda^k e^{-\lambda}}{k!}$, k = 0, 1, 2, …
- Mean of Y, $\mu_Y = \lambda$, since

$$E(Y) = \sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{k\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} =$$

$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Poisson Distribution - Variance

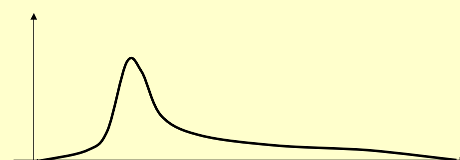
- Y~Poisson(λ), then P(Y=k) = $\frac{\lambda^k e^{-\lambda}}{k!}$, k = 0, 1, 2, …
- Variance of Y, $\sigma_Y = \lambda^{1/2}$, since

$$\sigma_Y^2 = Var(Y) = \sum_{k=0}^{\infty} (k-\lambda)^2 \frac{\lambda^k e^{-\lambda}}{k!} = \ldots = \lambda$$

- For example, suppose that Y denotes the number of blocked shots (arrivals) in a randomly sampled game for the UCLA Bruins men's basketball team. Then a Poisson distribution with mean=4 *may be* used to model Y .

Poisson Distribution - Example

- For example, suppose that Y denotes the number of blocked shots in a randomly sampled game for the UCLA Bruins men's basketball team. Poisson distribution with mean=4 may be used to model Y .



Poisson as an *approximation* to Binomial

- Suppose we have a sequence of Binomial(n, p_n) models, with $\lim(n\, p_n) \to \lambda$, as $n \to$ infinity.
- For each 0<=y<=n, if Y_n~ Binomial(n, p_n), then
  - $P(Y_n=y) = \binom{n}{y} p_n^{\,y} (1-p_n)^{n-y}$
  - But this converges to:

$$\binom{n}{y} p_n^{\,y} (1-p_n)^{n-y} \xrightarrow[n \times p_n \to \lambda]{\text{WHY?} \quad n \to \infty} \frac{\lambda^y e^{-\lambda}}{y!}$$

- Thus, Binomial(n, p_n) → Poisson(λ)

Poisson as an *approximation* to Binomial

- **Rule of thumb** is that approximation is good if:
  - **n>=100**
  - **p<=0.01**
  - **λ =n p <=20**
- Then, Binomial(n, p_n) → Poisson(λ)

Example using Poisson approx to Binomial

- Suppose P(defective chip) = 0.0001=10^{-4}. Find the probability that a lot of 25,000 chips has > 2 defective!
- Y~ Binomial(25,000, 0.0001), find P(Y>2). Note that Z~Poisson(λ =n p =25,000 x 0.0001=2.5)

$$P(Z>2)=1-P(Z\le 2)=1-\sum_{z=0}^{2}\frac{2.5^z}{z!}e^{-2.5}=$$

$$1-\left(\frac{2.5^0}{0!}e^{-2.5}+\frac{2.5^1}{1!}e^{-2.5}+\frac{2.5^2}{2!}e^{-2.5}\right)=0.456$$

Normal approximation to Binomial

- Suppose Y~Binomial(n, p)
- Then $Y=Y_1+Y_2+Y_3+\ldots+Y_n$, where
  - Y_k~Bernoulli(p) , $E(Y_k)=p$ & $Var(Y_k)=p(1-p)$ ➔
  - E(Y)=np & Var(Y)=np(1-p), SD(Y)= $(np(1-p))^{1/2}$
  - Standardize Y:
    - Z=(Y-np) / $(np(1-p))^{1/2}$
    - By CLT ➔ Z ~ N(0, 1). So, Y ~ N [np, $(np(1-p))^{1/2}$]
- **Normal Approx to Binomial is reasonable when np >=10 & n(1-p)>10** (p & (1-p) are NOT too small relative to n).

Normal approximation to Binomial – Example

- Roulette wheel investigation:
- Compute P(Y>=58), where Y~Binomial(100, 0.47) –
  - The proportion of the Binomial(100, 0.47) population having more than 58 reds (successes) out of 100 roulette spins (trials).
  - Since np=47>=10 & n(1-p)=53>10 Normal approx is justified.
- Z=(Y-np)/Sqrt(np(1-p)) = 58 – 100*0.47)/Sqrt(100*0.47*0.53)=2.2
- P(Y>=58) ⟵ ➔ P(Z>=2.2) = 0.0139
- True P(Y>=58) = 0.177, using SOCR (demo!)
- Binomial approx useful when no access to SOCR avail.

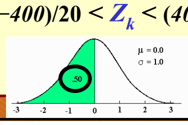
Roulette has 38 slots
18red 18black 2 neutral

Normal approximation to Poisson

- Let X_1~Poisson(λ) & X_2~Poisson(μ) ➔X_1+X_2~Poisson($\lambda+\mu$)
- Let $X_1, X_2, X_3, \ldots, X_k$ ~ Poisson(λ), and independent,
- $Y_k = X_1 + X_2 + \cdots + X_k$ ~ Poisson(kλ), $E(Y_k)=Var(Y_k)=k\lambda$.
- The random variables in the sum on the right are independent and each has the Poisson distribution with parameter λ.
- By CLT the distribution of the standardized variable $(Y_k - k\lambda) / (k\lambda)^{1/2}$ ➔ N(0, 1), as *k* increases to infinity.
- So, for $k\lambda$ >= 100, $Z_k = \{(Y_k - k\lambda) / (k\lambda)^{1/2}\}$ ~ N(0,1).
- ➔ Y_k ~ N($k\lambda$, $(k\lambda)^{1/2}$).

Normal approximation to Poisson – example

- Let X_1~Poisson(λ) & X_2~Poisson(μ) ➔X_1+X_2~Poisson($\lambda+\mu$)
- Let $X_1, X_2, X_3, \ldots, X_{200}$ ~ Poisson(2), and independent,
- $Y_k = X_1 + X_2 + \cdots + X_k$ ~ Poisson(400), $E(Y_k)=Var(Y_k)=400$.
- By CLT the distribution of the standardized variable $(Y_k - 400) / (400)^{1/2}$ ➔ N(0, 1), as *k* increases to infinity.
- $Z_k = (Y_k - 400) / 20$ ~ N(0,1)➔ Y_k ~ N(400, 400).
- P(2 < Y_k < 400) = (std'z 2 & 400) =
- P((2−400)/20 < Z_k < (400−400)/20) = P(-20< Z_k<0) = 0.5



Poisson or Normal approximation to Binomial?

- Poisson Approximation (Binomial(n, p_n) → Poisson(λ)):

$$\binom{n}{y} p_n^{\,y}(1-p_n)^{n-y} \xrightarrow[n\times p_n \longrightarrow \lambda]{\text{WHY?} \quad n\longrightarrow\infty} \frac{\lambda^y e^{-\lambda}}{y!}$$

  - n>=100 & p<=0.01 & λ =n p <=20
- Normal Approximation
  (Binomial(n, p) → *N* (np, $(np(1-p))^{1/2}$))
  - np >=10 & n(1-p)>10

Geometric, Hypergeometric, Negative Binomial

- X ~ Geometric(p), then the probability mass function is

Probability of first failure at x^{th} trial.

$$P(X=x)=(1-p)^x p; \quad E(X)=\frac{1-p}{p}; \quad Var(X)=\frac{1-p}{p^2}$$

- Ex: Stat dept purchases 40 light bulbs; 5 are defective.

  Select 5 components at random.

  Find: P(3^{rd} bulb used is the first that does not work) = ?

Geometric, Hypergeometric, Negative Binomial

- Hypergeometric – X~HyperGeom(n, M,N)

Total objects: N. Successes: M. Sample-size: n (without replacement). X = number of Successes in sample

$$E(X)=n\frac{M}{N}$$

$$Var(X)=\frac{N-n}{N-1}\times n\times\frac{M}{N}\times\frac{N-M}{N}$$

$$P(X=x)=\frac{\binom{M}{n}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Ex: 40 components in a lot; 3 components are defectives.

Select 5 components at random.

P(obtain one defective) = P(X=1) = ?

Hypergeometric Distribution & Binomial

- Binomial approximation to Hyperheometric
  - $\frac{n}{N}$ is small (usually < 0.1), then $\frac{k}{N} \approx p$

$$HyperGeom(x;N,n,k) \overset{approaches}{\Rightarrow} Bin(x;n,p)$$

Ex: 4,000 out of 10,000 residents are against a new tax. 15 residents are selected at random.

P(at most 7 favor the new tax) = ?

Geometric, Hypergeometric, Negative Binomial

- Negative binomial pmf [X ~ NegBin(r, p), if r=1 ➔ Geometric (p)] $\quad P(X=x)=(1-p)^x p$

Number of failures until the r^{th} success (negative, since number of successes (r) is fixed & number of trials (X) is random)

$$P(X=x)=\binom{x+r-1}{r-1}p^r(1-p)^x$$

$$E(X)=\frac{r(1-p)}{p}; \quad Var(X)=\frac{r(1-p)}{p^2}$$

Relation among Distributions