The Binomial Distribution

Data Science for Quality Management: Probability and Probability Distributions

with **Wendy Martin**

Learning objectives:

Describe the Binomial probability distribution

Calculate probabilities using the Binomial distribution

The Binomial Distribution

- The Binomial distribution relates to a discrete random variable (nominal data).

- The basis of this distribution is the Bernoulli process.

The Bernoulli Process

- Each trial or experiment has only two possible outcomes
- The P of any and all outcomes remains fixed over time
- The trials or experiments are statistically independent

The Binomial Formula

$$P(r\ in\ n\ trials) = \left[\frac{n!}{r!\,(n-r)!}\right][p^r][q^{n-r}]$$

where
p = probability of occurrence
q = 1-p = probability of failure
r = number of occurrences desired
n = number of trials

Binomial Example

- A vendor frequently ships 2 bad parts out of 10.
- Suppose the vendor ships our company 50 parts. If we tell them that at least 9 parts out of 10 must be good, and nothing in their manufacturing process has changed, what is the P that we will receive what we asked for?

Binomial Example

- p = 0.80, q = 0.20, r = 45, n = 50

$$P(45\ in\ 50) = \left[\frac{50!}{45!\,(50-45)!}\right][0.8^{45}][0.2^{5}]$$

= 0.02953

Binomial Example

- What if we wanted to know the probability of getting at least 9 out of 10 good parts in the shipment of 50? $P \geq 45$?
- We would sum the following:
- $P(45) + P(46) + P(47) + P(48) + P(49) + P(50)$

Probability Distributions

In R / Rstudio

> table.dist.binomial(n, p)

> pbinom( )

The Poisson Distribution

- This probability distribution is for discrete random variables which can take integer (whole) values (ordinal data).

Poisson Data Examples

- The number of parts produced during a 10 minute period
- The number of breakdowns per shift
- The number of failures per 100 cycles

The Poisson Formula

$$P(X) = \frac{\lambda^X}{X!} e^{-\lambda}$$

where
P(X) = probability exactly X occurrences
λ = Mean number of occurrences per time interval (or unit)
e = 2.71828

Poisson Example

- λ = 25 parts produced per hour
- X = 10 parts produced in one hour

$$P(10) = \frac{25^{10}}{10!} e^{-25}$$

= 0.0000365

Probability Distributions

In R / Rstudio

> table.dist.poisson(λ)

> ppois( )

Test for Poisson Distribution

In R / Rstudio

> poisson.dist.test( )

Sources

- Luftig, J. An Introduction to Statistical Process Control & Capability. Luftig & Associates, Inc. Farmington Hills, MI, 1982