Dot diagram

Time serries

Three basic methods of collecting data

hypothesis testing

hypothesis

mechanistic model

empirical model

sample and population

Probability models

Statistics and statistic

Statistical inference

Regression variable

The field in statistic and probability:

the collection, presentation, analysis and use of data to make decisions, solve problems.

quantifying the risks involed in decisions made every day.

1. A sample of two items is selected without replacement

from a batch. Describe the (ordered) sample space for

each of the following batches:

(a) The batch contains the items {*a*, *b*, *c*, *d*}.

(b) The batch contains the items {*a*, *b*, *c*, *d*, *e*, *f*, *g*}.

(c) The batch contains 4 defective items and 20 good items.

(d) The batch contains 1 defective item and 20 good items.

2. A sample of two printed circuit boards is selected

without replacement from a batch. Describe the (ordered)

sample space for each of the following batches:

(a) The batch contains 90 boards that are not defective, 8

boards with minor defects, and 2 boards with major

defects.

(b) The batch contains 90 boards that are not defective, 8

boards with minor defects, and 1 board with major

defects.

1. **A month of the year is chosen at random. What’s is the probability of choosing a April or May?**
2. **A school survey found that 7 out of 30 workers walk to company. If four workers are selected at random without replacement, what is the probability that all four walk to company?**
3. A batch of 60 containers for frozen orange juice contains 7 that are defective. Two are selected, at random, without replacement, from the batch. Let A and B denote the events that the first and second container selected is defective, respectively.  
   Are A and B independent events?
4. Suppose that P(B|A) = 0.8, P(A) = 0.2 and P(B) = 0.4. Find P(A|B).
5. Suppose we take all the different families with exactly 4 children. The experiment consists in asking them the sex (or genders) of the first and the second ,…children. Write down the sample space. Let us write 'B' for boy and 'G' for girl.



1. A discrete random variable X with possible values 0, 1,2,3,4 and the probability mass function of X is f( x )=  . Find P(X<2) and P(X=5);
2. Determine the cumulative distribution function for the variable variable X with possible values 0, 1,2,3,4 and the probability mass function of X is f( x ) =  .
3. Determine the mean and variance of the random variable X with possible values 0, 1,2,3,4 and the probability mass function of X is f( x )=  .
4. Let X denote the number of bits received in error in a digital communication channel, and assume that X is a binomial random variable with p = 0.025. If 25 bits are transmitted, find P( X=5 ); and the mean of X.
5. Suppose Y has a hypergeometric distribution with N=30, n=5 and K=8. Find P(Y>2) and the mean of Y.
6. Suppose that *X* is a negative binomial random variable with *p* = 0.3 and *r* =4. Determine the following:

P(X=20) and V(X).

1. The number of telephone calls that arrive at a phone exchange is often modeled as a Poisson random variable. Assume that on the average there are 8 calls per hour.

(a) What is the probability that there are exactly 3 calls in one

hour?

(b) What is the probability that there are 3 or less calls in one

hour?

(c) What is the probability that there are exactly 13 calls in

two hours?

(d) What is the probability that there are exactly 5 calls in

30 minutes?

1. If X is a continuous random variable with probability density function f( x ) = x/3,  <x < 3.Find the variance of X; P(X=2); P(X<2)
2. Assume X is normally distributed with a mean of 12 and a standard deviation of 1.05. Determine the value for x that holds:  
   P( X >x) = 0.5. Let P(Z<0)= 0 .5; Z is a standard normal random variable.
3. Suppose *X*has a continuous uniform distribution over the interval [-2,5]. Determine the mean, variance, and standard deviation of *X.*
4. Let X be the continuous random variable with cumulative distribution function F( x ) = 1 – e-k.x , x>0; find the probability density function of X.
5. **The time between the arrival of electronzic messages at your computer is exponentially** distributed with a mean of two hours. What is the probability that you do not receive a message during a two-hour period?
6. The time between arrivals of small aircraft at a county airport is exponentially distributed with a mean of one hour. What is the probability that more than three aircraft arrive within an hour?
7. Assume that the number of asbestos particles in a squared meter of dust on a surface follows a Poisson distribution with a mean of 2000. If a squared meter of dust is analyzed, what is the probability that less than 950 particles are found?  
   Let P( Z<-4) = 0; P(Z <5) = 1; P( Z<0) = 0.5; Z is a standard normal random variable.
8. An article in the ASCE Journal of Energy Engineering (1999, Vol. 125, pp. 59–75) describes a study of the thermal inertia properties of autoclaved aerated concrete used as a building material. Five samples of the material were tested in a structure, and the average interior temperature (°C) reported was as follows: 23.01, 22.22, 22.04, 22.62, 23.01 and 22.59.

Find the sample mean, standard deviation, range, mode and variance.

1. The compressive strength of concrete is normally distributed with =1150 psi and =35. A random sample of n=35 specimens will have a sample mean diameter that falls in the interval from 1141 psi to 1230 psi. What is the standard error of the sample mean? How large must be the random sample be if we want the standard error of the sample mean to be 2.4?



1. A random sample of size n1=36 is selected from a normal population with a mean of 35 and a standard deviation of 2. A second random sample of size n2=64 is taken from another normal population with a mean of 40 and standard deviation 3. Let be the two sample means. Find P(<1.5).
2. Data on oxide thickness of semiconductors are as follows: 135, 141, 126, 129, 131, 146, 128, 120, 141, 143. The point estimate of the mean, the standard deviation and the variance of oxide thickness.
3. A civil engineer is analyzing the compressive strength of concrete. Compressive strength is normally distributed with = 1000(psi)2. A random sample of 12 specimens has a mean compressive strength of  *=*3250 psi.

(a) Construct a 95% two-sided confidence interval on mean compressive strength. To estimate the compressive strength with an error that is less than 15 psi at 99% confidence. What sample size is required?

(b) Construct a 99% two-sided confidence interval on mean compressive strength.

1. The brightness of a television picture tube can be evaluated by measuring the amount of current required to achieve a particular brightness level. A sample of 10 tubes results in * = 317.2* and *s* = 15.7. Find (in microamps) a 99% confidence interval on mean current required.
2. A rivet is to be inserted into a hole. A random sample of *n* =15 parts is selected, and the hole diameter is measured. The sample standard deviation of the hole diameter measurements is *s* =0.008 millimeters. Construct a 99% lower confidence bound for .
3. A manufacturer of electronic calculators is interested in estimating the fraction of defective units produced. A random sample of 950 calculators contains 15 defectives. Compute a 99% upper-confidence bound on the fraction defective. How large must the sample be if we wish to be at least 95% confident that the error in estimating *p* is less than 0.02, regardless of the true value of *p*?
4. Which correlation coefficient represents the strongest association between the *X* and *Y* variables?
5. *r* = -0.80
6. *r* = -0.30
7. *r* = +0.50
8. *r* = -0.70
9. If = 3.8, = 21 and = 2.2, then the y-intercept of the least-squares regression line is?
10. Given the equation of a regression line is = 2.7x + 3.2, what is the best predicted value for y given x =-2.2. Assume that the variables x and y have a significant correlation.
11. Suppose data are obtained from 22 pairs of (x,y) and the sample correlation coefficient is 0.55. Test the hypothesis that H0: ρ=0 against H1: ρ≠0 with α =0.05.
12. A manufacturer produces crankshafts for an automobile engine. The wear of the crankshaft after 100,000 miles (0.0001 inch) is of interest because it is likely to have an impact on warranty claims. A random sample of *n* =15 shafts is tested and 2.78. It is known that =0.9 and that wear is normally distributed. Test *H*0: = 3 versus H1:3 using α = 0.05.
13. A 1992 article in the *Journal of the American Medical Association* (“A Critical Appraisal of 98.6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wundrlich”) reported body temperature, gender, and heart rate for a number of subjects. The body emperatures for 25 female subjects follow: 97.8, 97.2, 97.4, 97.6, 97.8, 97.9, 98.0, 98.0, 98.0, 98.1, 98.2, 98.3, 98.3, 98.4, 98.4, 98.4, 98.5, 98.6, 98.6, 98.7, 98.8, 8.8, 98.9, 98.9, and 99.0. Test the hypotheses *H*0:  = 98.6 versus , H 1 : 98.6 using α = 0.05; Test the hypothesis that  = 0.10 against an alternative specifying that   0.10, using α = 0.01.
14. The fraction of defective integrated circuits produced in a photolithography process is being studied. A random sample of 300 circuits is tested, revealing 13 defectives. Test *H*0: *p =* 0.05 versus *H*1: *p*  0.05. Use α = 0.05.
15. Based on the data from six students, the regression equation relating number of hours of preparation (x) and test score (y) is ŷ = 6.3 + 0.97x. The same data yield r = 0.425 and ȳ = 6.5. What is the best predicted test score for a student who spent 3 hours preparing for the test?
16. Based on the data from six students, the regression equation relating number of hours of preparation (x) and test score (y) is ŷ = 6.3 + 0.97x. The same data yield r = 0.825 and ȳ = 6.5. What is the best predicted test score for a student who spent 3 hours preparing for the test?
17. The following data consists of test scores and hours of preparation for 5 randomly selected students in a class. For example, the student who spent 6 hours preparing for the test had a score of 73.

Hours of Preparation: 5 2 9 6 10  
Test Score: 64 48 72 73 82

1. Find the equation of the regression line for this data. Round values to three decimal places.
2. Compute the correlation coefficient r.
3. What is the best predicted test score for a student who spent 8 hours preparing for the test?
4. Given a sample with r = 0.8, n=20 and α = 0.01, determine the standardized test statistic t0 necessary to test the claim ρ = 0.