Ivan Smirnov

Statistics Study Sheet

:: Key Terms ::
- statistic – numerical descriptive of sample
- parameter – numerical descriptive of population
- expected value – what we expect the mean to be
- confidence interval – the # in $\mu \pm$ # that a certain percentage of data is certain to fall within this range
- margin of error – general formula (invNorm * (0.5 + {CONFIDENCE INTERVAL / 2}) x (μ /√ n)

:: Symbols ::

n – number of trials
p – probability of success
q – probability of failure (used in binomial probabilities)
r – # of successes (used in binomial probabilities)
$\bar{x}$ – mean (statistic)
μ – mean (parameter)
s^2 – variance (statistic)
σ^2 – variance (parameter)
s – standard deviation (statistic)
σ – standard deviation (parameter)
$\hat{p}$ – proportion (statistic)
p – proportion (parameter)

:: Methods ::

A. one list of data – use List → 1-var stats. You can get the mean, median, mode, quarters, and std-dev
B. two lists of related data (x,y) – use 1-var stats, but pass L_1 , L_2 as the argument
C. Binomial Distribution – use binomcdf(n,p,r)
    $\sqrt{(npq)} = \sigma$ and $\mu = np$
D. Normal distribution $\mu=0$, and $\sigma=1$
    use z-values to manipulate data
    invNorm(z-score, μ, σ)= x-value
E. Distribution of means. If not normal or unknown, n > 30 must be true
    new σ = (regular σ)/(√n)
F. Matrix (a chart) use X^2 test (chi-squared test)
G. Two lists – use ANOVA test

:: Problem Solving Checklist ::
1. Which formula should you use? (Consult study sheets)
2. are you finding t-value or z-score?
3. Is the test one tailed or two tailed?

:: Syntax Help ::

normalpdf(#)-gives the probability of that EXACT number occurring
normalcdf(lower limit, upper limit, mean, std-dev)
invNorm(are on the left, mean, sigma)
binompdf(n,p,r) – probability of r occurring
binomcdf(n,p,r) – sum of probabilities from 0 to r
tcdf(lower, upper, degrees of freedom)
invT(area, degrees of freedom)

*NOTE – Don't blindly trust the calculator, always check to make sure the answer makes sense!

:: Extra ::

Chebyshev formula gives us % of data within k std-devs IFF $k>1$
% data = $1 - (1/k^2)$
Note – a number greater than 2.5std-devs away from the mean is a rare event, an outlier

Normal distribution
68% within 1σ
95% within 2σ
99.7 within 3σ