

# DATA1001 Topic 7 Flashcards

## Chance Variability (The Box Model)

Abyan Majid

Question	My answer									
Recite gambler’s fallacy and the law of averages!	Gambler’s fallacy is the mistake of assuming that the chance of observing a particular event changes over time  Law of averages states that the observed proportion of an event will eventually correspond to the expected proportion in the long run.									
Identify 3 pieces of information you need to know in order to use the box model?	1. All distinct numbers (“tickets”) in the box 2. The amount of each unique “ticket” in the box (accounting for repetitions) 3. Number of draws from the box									
We can think of the box as a ..., and whatever is drawn from it creates the ...	Summary of the population, sample									
Recite the formula for chance error (CE)!	$chance\ error = observed\ value + expected\ value$									
How to compute: (1) SD of the box? (2) SD of a binary box (shortcut alternate)?	(1) $SD_{pop} = RMS(gaps)$ (2) $SD_{pop} = max - min \sqrt{prop.\ of\ max \times prop.\ of\ min}$									
For the SUM of random draws from a box model with replacement, how do we find: (3) Observed value (OV)? (4) Expected value ()? (5) Standard error?	(1) $observed\ value = expected\ value + chance\ error$ (2) $expected\ value = no.\ of\ draws \times mean\ of\ box$ (3) $standard\ error = \sqrt{no.\ of\ draws} \times SD\ of\ box$ <table><tr><td></td><td>EV</td><td>SE</td></tr><tr><td>Sum of the Sample</td><td><math>n \times mean</math></td><td><math>\sqrt{n} \times SD</math></td></tr><tr><td>Mean of the Sample</td><td>mean</td><td><math>SD / \sqrt{n}</math></td></tr></table>		EV	SE	Sum of the Sample	$n \times mean$	$\sqrt{n} \times SD$	Mean of the Sample	mean	$SD / \sqrt{n}$
	EV	SE								
Sum of the Sample	$n \times mean$	$\sqrt{n} \times SD$								
Mean of the Sample	mean	$SD / \sqrt{n}$								
For the MEAN of random draws from a box model with replacement, how do we find: (1) Observed value? (2) Expected value? (3) Standard error?	(1) $observed\ value = expected\ value + chance\ error$ (2) $expected\ value = mean\ of\ box$ (3) $standard\ error = \frac{population\ SD}{\sqrt{no.\ of\ draws}}$									

	<table><tr><td></td><td>EV</td><td>SE</td></tr><tr><td>Sum of the Sample</td><td><math>n \times \text{mean}</math></td><td><math>\sqrt{n} \times \text{SD}</math></td></tr><tr><td>Mean of the Sample</td><td>mean</td><td><math>\text{SD} / \sqrt{n}</math></td></tr></table>		EV	SE	Sum of the Sample	$n \times \text{mean}$	$\sqrt{n} \times \text{SD}$	Mean of the Sample	mean	$\text{SD} / \sqrt{n}$
	EV	SE								
Sum of the Sample	$n \times \text{mean}$	$\sqrt{n} \times \text{SD}$								
Mean of the Sample	mean	$\text{SD} / \sqrt{n}$								
Given a sum of random draws from a box model with replacement, is it possible to pass EV and SE as parameters to a normal model?	Yes									
Identify 3 types of histograms and what each represents!	(1) Data histogram: It represents frequency of data by area (2) Probability histogram: It represents chance by area (3) Simulation histogram: It represents chance by area for a simulation of a chance process									
What does EV and SE represent in a probability histogram?	EV represents the center of the horizontal axis. Meanwhile, SE represents the spread of the horizontal axis.									
Recite the central limit theorem (CLT) and any conventional conditions deemed sufficient for CLT to apply!	If the sample size for the sum (or mean) is sufficiently large then the probability histogram for the sum (or mean) will follow the normal curve  By convention, if the the sample size/number of draws is large enough ( $n > 30$ ), then we usually consider CLT to apply.									