University Introductory Statistics · Practice Worksheet · tutorinmontreal.ca

Sampling Distributions and the Central Limit Theorem

Name: ______________________________ Date: ______________

Concepts

- Random Samples, Statistics and Unbiased Estimators (theory)
- Mean and Variance of a Sum of Independent Variables (theory)
- The Sampling Distribution of the Sample Mean
- The Central Limit Theorem
- The Sampling Distribution of a Sample Proportion

Random Samples, Statistics and Unbiased Estimators

1. A population consists of the three values 1, 4 and 7. Its mean is μ and its variance is $\sigma^2 = \frac{1}{3}\sum(x-\mu)^2$, the average squared deviation over the whole population. A random sample of size $n = 2$ is drawn *with replacement*, so the two draws are independent and each of the 9 ordered samples is equally likely.

a) Find μ and σ^2.

b) List the 9 samples with the sample mean $\bar{x}$ of each, and give the sampling distribution of $\bar{X}$ as a table. Show that $E[\bar{X}] = \mu$.

c) For each sample compute the sample variance $s^2 = \dfrac{\sum(x_i - \bar{x})^2}{n-1}$. Show that the average of s^2 over the 9 samples equals σ^2, and say what would happen to that average if the divisor were n instead of $n-1$.

Mean and Variance of a Sum of Independent Variables

2. At a small ferry terminal, the time a car spends boarding, B, has mean 18 minutes and standard deviation 5 minutes, and the crossing time C has mean 40 minutes and standard deviation 12 minutes. B and C are independent.

a) Find the mean and the standard deviation of the total time $B + C$.

b) Find the mean and the standard deviation of the difference $C - B$.

The Sampling Distribution of the Sample Mean

3. The mass of jam in a jar from a filling line is normally distributed with mean 250 g and standard deviation 20 g; we write this $N(250,\, 20^2)$, where the second argument of $N(\mu, \sigma^2)$ is the *variance*. Use $P(Z < -0.25) = 0.4013$ and $P(Z < -1.00) = 0.1587$.

a) Find the probability that one randomly chosen jar holds less than 245 g.

b) A random sample of 16 jars is taken. Give the mean and the standard deviation of the sample mean $\bar{X}$, and the distribution of $\bar{X}$.

c) Find the probability that the sample mean of the 16 jars is less than 245 g.

4. Resting heart rates of adults in a large fitness program follow $N(70,\ 10^2)$, in beats per minute, where the second argument is the variance. Use $z_{0.025} = 1.960$, the value with $P(Z > 1.960) = 0.025$.

a) Find the interval, symmetric about 70, that contains the heart rate of one randomly chosen participant with probability 0.95.

b) Find the interval, symmetric about 70, that contains the mean heart rate $\bar{X}$ of a random sample of 25 participants with probability 0.95.

The Central Limit Theorem

5. The duration of a call to a software help desk has mean 8 minutes and standard deviation 6 minutes, and its distribution is moderately skewed to the right: most calls are short, a few are much longer. A random sample of 36 calls is taken, independently.

The **central limit theorem** says: if $X_1, \ldots, X_n$ are independent observations from one population with mean μ and finite standard deviation σ, then for n large enough the distribution of $\bar{X}$ is approximately $N(\mu,\ \sigma^2/n)$ (second argument the variance), whatever the shape of the population; the more skewed the population, the larger n must be, and $n \geq 30$ is a common working guide for a moderately skewed one. Use $P(Z < 1.50) = 0.9332$.

a) Give the approximate distribution of the sample mean duration $\bar{X}$, and say which condition of the theorem the sample size is there to satisfy.

b) Find the approximate probability that the 36 calls average more than 9.5 minutes.

c) Can the same method give the probability that *one* call lasts more than 9.5 minutes? Answer in one sentence.

6. The number of weaving flaws on a bolt of fabric has mean 2.4 and standard deviation 1.5; it is a count, $0, 1, 2, \ldots$, and is not normally distributed. An inspector examines a random sample of 100 bolts, whose flaw counts are independent. Use $P(Z < -2.00) = 0.0228$, $P(Z < -1.00) = 0.1587$ and $P(Z < 2.00) = 0.9772$.

a) Explain why the sample mean number of flaws $\bar{X}$ may be treated as normal, and give its mean and standard deviation.

b) Find the approximate probability that $\bar{X} < 2.1$.

c) Find the approximate probability that $2.25 < \bar{X} < 2.7$.

The Sampling Distribution of a Sample Proportion

7. Of the seeds in a large seed lot, a proportion $p = 0.20$ fail to germinate. A random sample of 400 seeds is planted, and $\hat{p}$ is the proportion that fail. The normal model for $\hat{p}$ may be used when $np \geq 10$ and $n(1-p) \geq 10$; use no continuity correction. Use $P(Z < 2.00) = 0.9772$ and $P(Z < 1.50) = 0.9332$.

a) Check the condition, and give the mean and standard deviation of $\hat{p}$.

b) Find the approximate probability that more than 24% of the sample fail.

c) Find the approximate probability that between 17% and 23% of the sample fail.

8. In a city, 64% of library cardholders borrow at least one e-book a year. The normal model for a sample proportion $\hat{p}$ may be used when $np \geq 10$ and $n(1-p) \geq 10$; use no continuity correction. Use $P(Z < -1.50) = 0.0668$ and $P(Z < -3.00) = 0.0013$.

a) For a random sample of 144 cardholders, find the standard deviation of $\hat{p}$ and the approximate probability that fewer than 58% of the sample borrow e-books.

b) Repeat for a random sample of 576 cardholders.

Synthesis — drawing on several topics in this unit

9. A delivery hub loads 100 parcels onto a van. Parcel masses vary independently, with mean 2.5 kg and standard deviation 1.2 kg, and are skewed to the right. The van's payload limit is 268 kg. Use $P(Z < 1.50) = 0.9332$ and $z_{0.025} = 1.960$, the value with $P(Z > 1.960) = 0.025$.

a) Find the mean and standard deviation of the total mass T of the 100 parcels.

b) Explain why T is approximately normal, and find the approximate probability that the load exceeds the payload limit.

c) Find the load that is exceeded with probability only 0.025.