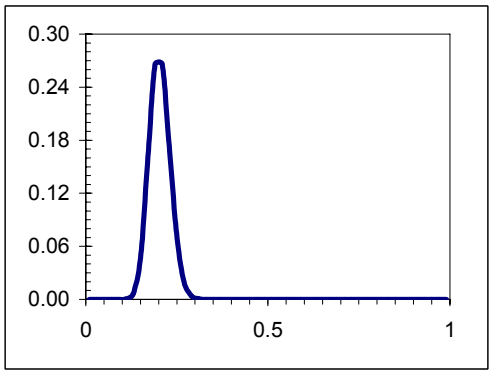
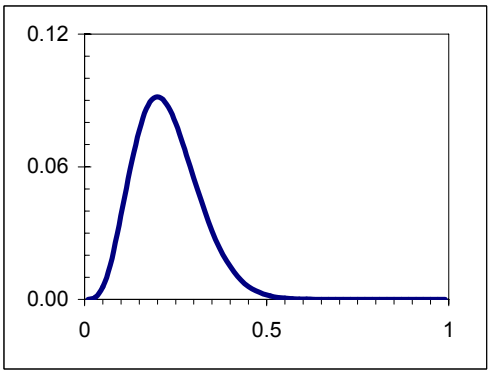
Biostatistics **Sample Exam 2**

1. A large population contains an unknown proportion (p) of black marbles. A sample of n=200 drawn scientifically from the population contained x=40 black marbles. Which picture shows the posterior distribution of the population proportion p? Justify your answer.

A. Looks like this (m = .2 and s = .028)



B. Looks like this (m= .2 and s = .089)



2. An investigator wants to determine the proportion (p) of retirees who chose not to fill a prescription last year because it was too expensive. In a scientific sample of n=525, x = 98 respondents said that they had done this. State in words the posterior distribution of p. Obtain a 95% credible interval for p. What (approximately) is the probability that p exceeds 0.25 ?

3. A prospective study recruited 10,000 smokers and 10,000 non-smokers aged 30 to 39 and followed them for 20 years. The cases of throat cancer are shown in this table.

	no Cancer	Cancer	%
Smokers	9800	200	2.0%
Non-Smokers	9950	50	0.5%

What is the relative risk of throat cancer for smokers vs. non-smokers?

4. Two hundred forty dieters volunteered for a study of chromium picolinate, a fat-reducing dietary supplement. They were randomly assigned to receive placebo or chromium picolinate. One side effect is reduction in iron, a key component of hemoglobin. Here are the data on percents of subjects with lower iron after 8 weeks of treatment.

	Placebo	Chromium Picolinate
n	120	120
% with reduced iron	16%	29%

- What are the mean difference and the standard error of the difference (SED) between the percents?
- Obtain a 95% credible interval for the difference.
- Is the difference statistically significant?

5. In a study to compare osteoporosis rates for men and women over the age of 70 it was observed that 6.9% of 25000 men and 67% of 27000 women had osteoporosis. The difference is 60.1 percentage points and the credible interval is 59.4 to 60.7. Is the difference significant?

6. In a randomized experiment, 400 kids brushed with baking powder and 400 brushed with toothpaste. 52 of the baking powder kids (13%) got cavities and 40 of the toothpaste kids (10%) got cavities. The difference is 3 percentage points. Obtain a 95% credible interval for the true difference. Is the difference significant?

7. An economic survey of a sample of 225 US wage earners showed an average of $\bar{x}$ =\$23.50 was spent per week eating out. The standard deviation of the sample was reported to be s = \$12.00. State the posterior distribution and obtain a 95% credible interval on the mean (μ) of all wage earners.

8. A random roadside survey of 481 males and 138 females found that 77 males and 16 females had detectable amounts of alcohol by a breathalyzer test. State the posterior distribution of the difference. Obtain a 95% credible interval for the difference. Is the difference statistically significant?

9. R. M. Lyle, reported a study in which healthy men aged 45 to 65 received either a calcium supplement or a placebo for 12 weeks. He reported, "The calcium group had lower blood pressure (p=.008) compared with the placebo group." (Note: blood pressure is measured in millimeters of mercury, abbreviated mm Hg.)

Which of the following sets of data is consistent with Lyle's statement? Why?

A: Difference = 10 mm Hg with 95% credible interval 2.4 to 17.6.
B: Difference = 20 mm Hg with 95% credible interval -5 to 45.

10. A sample drawn from a box of numbers with a fairly normal distribution has sample mean $\bar{x}$ = 16.5 and sample standard deviation s = 8.8. State the approximate posterior distribution of the box average (μ_{box})

a) if n=400 b) if n=36

11. An unknown quantity, which we will call η, has an approximately t(9) distribution with μ = 3.1 and σ = 0.6 Find the 95% confidence interval for the unknown quantity.

12. One hundred male alcoholics suffering from secondary hypertension participated in a study to determine the efficacy of a new antihypertensive agent. The men were assigned at random to either the control group or the treatment group. Men in the control group received a placebo. Statistics for arterial pressure at 30 days post treatment for the 97 subjects who completed the study are shown in this Table.

Hypertension Study	Placebo	Treatment
n	22	23
mean ($\bar{x}$)	127.1	99.0
standard deviation (s)	24.08	8.81

State the approximate posterior distribution of the difference (Δ) between the means of the two populations.

The posterior distribution is: Approximately t() with _____

95% Confidence interval: _____ to _____

Formulas

$$sem = \frac{s}{\sqrt{n}}$$

$$sep = \sqrt{\frac{\hat{p}\cdot(1-\hat{p})}{n}}$$

$$sed = \sqrt{(se_1)^2 + (se_2)^2}$$

$$Satterthwaite's\ df = \frac{(sed)^4}{\frac{(sem_1)^4}{df_1} + \frac{(sem_2)^4}{df_2}}$$