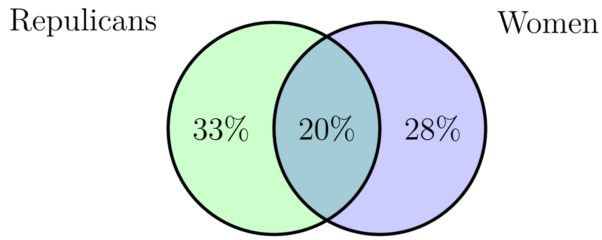
Statistics - Math 121 **Midterm 2 Practice Exam**

This is a practice exam for Midterm 2. If you have any questions, send me an e-mail!

1. In one congressional district, 53% of the voters support the Republican candidate, 48% of the voters are women, and 20% of the voters are women who support the Republican candidate.

   (a) Make a Venn diagram for this situation.

   Solution:

   (b) What percent of the women in the district support the Republican candidate?

   Solution: Notice the keyword: "of". This isn't asking for the percent **of people** who are women and support the Republican candidate. This is asking for the percent **of women** who support the Republican. That's not the same thing! The percent of women is the conditional probability that someone supports the Republican given that they are a woman:

$$P(\text{Republican} \mid \text{Woman}) = \frac{P(\text{Both})}{P(\text{Woman})} = \frac{20\%}{48\%} = 41.67\%.$$

   (c) Are gender and political preference independent?

   Solution: They are **not independent** because women have a lower chance of supporting the Republican candidate. Only 41.67% of women support the Republican vs. 53% of voters overall. *If you want another good conditional probability practice problem, find the percent of men who support the Republican candidate.*

2. The average weight of an American adult is 170 lbs. with a population standard deviation of about 40 lbs.

   (a) Describe the shape, center, and spread of the sampling distribution for the average weight $\bar{x}$ of a group of 25 random adults.

   Solution:

   Shape The sample size of 25 is big enough that the distribution will be roughly normal.

   Center The theoretical mean for $\bar{x}$ is the same: $\mu_{\bar{x}} = 170$ lbs.

   Spread The standard deviation for $\bar{x}$ is smaller:

   $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{40}{\sqrt{25}} = 8 \text{ lbs.}$$

   (b) Find the probability that average weight of the group is over 180 lbs.

   Solution: This is a bell curve problem. We have a normal distribution with $\mu_{\bar{x}} = 170$ and $\sigma_{\bar{x}} = 8$, so you can use the Probability Distributions app to find:

   $$P(\bar{x} \geq 180) = 0.10565 = 10.6\%.$$

3. A burger restaurant is having a special \$1 burger night. Research has shown that 20% of customers will just purchase a burger resulting in a loss to the restaurant of \$0.75, 30% of customers will also purchase fries with their burger resulting in a profit of \$0.50, and the rest opt for the full meal of a burger fries and a drink resulting in a profit of \$1.50. What is the expected (theoretical average) profit per customer? Hint: *The info above is also in this table:*

Outcome	−\$0.75	\$0.50	\$1.50
Probability	0.20	0.30	0.50

Solution: The expected value is the **weighted average**. So the average profit is

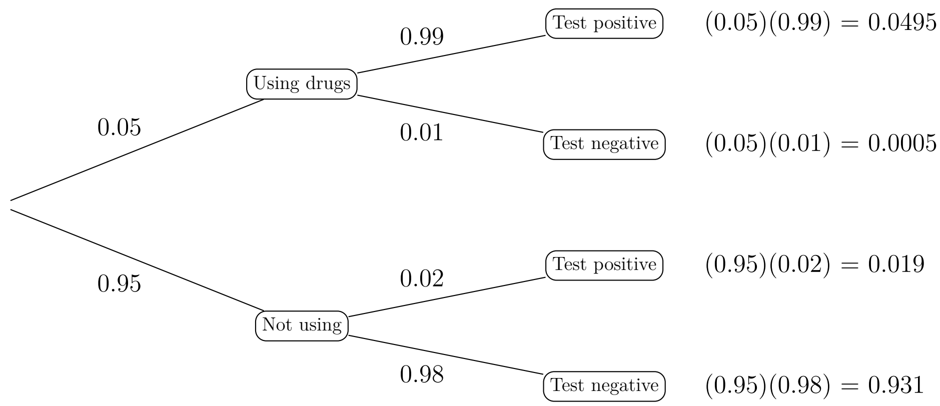
$$0.20(-0.75) + 0.30(0.50) + 0.5(1.50) = \$0.75$$

per customer, on average.

4. A company requires job applicants to take a drug test. The drug test is 99% accurate at detecting if someone is currently using illegal drugs. The test is also 98% accurate at recognizing if someone is not currently using illegal drugs (so it has a 2% false positive rate). Suppose that 5% of the applicants are using illegal drugs.

   (a) Draw and label a tree diagram for this situation.

   Solution:



   (b) What percent of applicants will test positive for illegal drug usage?

   Solution: There are two groups of people who will test postive: drug users who test positive and also some false positives. From the tree-diagram (multiplying weights along branches) we found that 4.95% of people are drug users who test positive, and 1.9% are non-drug users who get a false positive. So the total who test positive is 6.85%.

   (c) Given that a job applicant tests positive, what is the conditional probability that they have been using illegal drugs.

   Solution:

$$P(\text{Use drugs} \,|\, \text{Test positive}) = \frac{P(\text{Both})}{P(\text{Test positive})} = \frac{4.95}{6.85} = 72.26\%.$$

5. What does the Law of Large Numbers say about gambling at a casino?

   Solution: The Law of Large Numbers says that the more you gamble, the closer your average winnings tends to get to the theoretical average. Since the theoretical average winnings at a casino is less than it costs to play, you should expect to lose money if you play a lot.

6. Suppose you want to estimate the average resting heart rate of college athletes. You select a random sample of 30 college athletes, and find a sample mean of 60.3 beats per minute. You know that the standard deviation for heart rate is $\sigma = 6$.

   (a) Find a 95% confidence interval for the heart rate of college athletes.

   Solution: IF we know the population standard deviation σ, then we can use the normal distribution confidence interval formula

   $$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}.$$

   We know that $z^* = 1.96$ when you want a 95% confidence level (from the table on the formula sheet). Therefore, the confidence interval is:

   $$60.3 \pm 2.15$$

   which goes from 58.15 to 62.45 beats per minute.

   (b) What is the margin of error of the confidence interval?

   Solution: The margin of error is 2.15 beats per minute.

   (c) Explain what we are 95% confident about.

   Solution: Since this is a random sample from the population of all college athletes, we can be 95% sure that our confidence interval contains the true average resting heart rate for that population.

   (d) If you didn't know σ, but you calculated a sample standard deviation $s = 6.3$ beats per minute, then how would you calculate the confidence interval? What formula would you use, and what numbers would you plug into the formula?

   Solution: You would use the t-distribution confidence interval formula:

   $$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

   where $\bar{x} = 60.3$, $s = 6.3$, $n = 30$, and $t^* = 2.045$.

7. Researchers interested in lead exposure due to car exhaust sampled the blood of 52 police officers subjected to constant inhalation of automobile exhaust fumes while working traffic enforcement in a primarily urban environment. The blood samples of these officers had an average lead concentration of 124.32 micrograms per liter (μg/l) and a SD of 37.74 μg/l; a previous study of individuals from a nearby suburb, with no history of exposure, found an average blood level concentration of 35 μg/l.

   (a) Write down the hypotheses that would be appropriate for testing if the police officers appear to have been exposed to a different concentration of lead.

   Solution: The correct hypotheses are:

   - $H_0 : \mu_{\text{police}} = 35$
   - $H_A : \mu_{\text{police}} > 35$.

   I'd be okay if you used a two-sided alternative here too. But since we think that police officers work in a area with a lot of automobile exhaust fumes, I think a 1-sided alternative is appropriate.

   (b) Test the hypothesis that the downtown police officers have a higher lead exposure than the group in the previous study. Interpret your results in context.

   Solution: This is a t-Test. The t-value is:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{124.32 - 35}{37.74/\sqrt{52}} = \frac{89.32}{5.23} = 17.1$$

   That's a pretty big t-value! But we still need to find the p-value. Use the Probability Distribution app. I got:

$$p = 0$$

   which just means that the result is extremely significant. It is almost impossible for this to be a random fluke, so we reject the null hypothesis and conclude that police officers really are exposed to more lead than other people.

Formulas & Tables

Standardized normal data

$$z = \frac{x - \mu}{\sigma} \text{ or } \frac{\text{location} - \text{middle}}{\text{std. dev.}}$$

Standard deviation of sample mean

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

Basic Probability Rules

1. Addition Rule $\quad P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$
2. Multiplication Rule $\quad P(A \text{ and } B) = P(A)P(B|A)$

Conditional Probability

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

Confidence Intervals for μ

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} \qquad \text{or} \qquad \bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

Use the t-Distribution table (see next page) if you don't know σ to find the critical t-value, t^*. Remember that $\text{df} = n - 1$. Otherwise, use the table below to find z^*.

Common Confidence Levels

Level	90%	95%	99%	99.9%
z^*	1.645	1.960	2.576	3.291

Hypothesis Tests for μ

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \qquad \textit{or} \qquad t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Only use z-values and the normal distribution if you know σ. Otherwise use the t-Distribution to find the p-values. Remember that $\text{df} = n - 1$.

t-Distribution Critical Values

Confidence Level	60%	80%	90%	95%	98%	99%	99.8%	99.9%
df								
1	1.376	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.061	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.978	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.941	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.920	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.906	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.896	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.889	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.883	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.879	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.876	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.873	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.870	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.868	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.866	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.865	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.863	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.862	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.861	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.860	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.859	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.858	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.858	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.857	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.856	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.856	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.855	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.855	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.854	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.854	1.310	1.697	2.042	2.457	2.750	3.385	3.646
31	0.853	1.309	1.696	2.040	2.453	2.744	3.375	3.633
32	0.853	1.309	1.694	2.037	2.449	2.738	3.365	3.622
33	0.853	1.308	1.692	2.035	2.445	2.733	3.356	3.611
34	0.852	1.307	1.691	2.032	2.441	2.728	3.348	3.601
35	0.852	1.306	1.690	2.030	2.438	2.724	3.340	3.591
36	0.852	1.306	1.688	2.028	2.434	2.719	3.333	3.582
37	0.851	1.305	1.687	2.026	2.431	2.715	3.326	3.574
38	0.851	1.304	1.686	2.024	2.429	2.712	3.319	3.566
39	0.851	1.304	1.685	2.023	2.426	2.708	3.313	3.558
40	0.851	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	0.849	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	0.848	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.846	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.845	1.290	1.660	1.984	2.364	2.626	3.174	3.390
500	0.842	1.283	1.648	1.965	2.334	2.586	3.107	3.310
z^*	0.842	1.282	1.645	1.960	2.326	2.576	3.090	3.291
One-sided p-value	0.2	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
Two-sided p-value	0.4	0.2	0.1	0.05	0.02	0.01	0.002	0.001