# Public Policy 529
# Quiz #2

Student ID number (8-digits): ________________________

1. City government officials want to assess the effectiveness of a water conservation program. In participating households, mean water consumption was 400 gallons/day ($n$=61; $s$=50). In the control group, mean water consumption was 415 gallons/day ($n$=63; $s$=60). Assume samples were randomly selected.

   (a) Calculate the standard error of the difference and explain why the formula that you used for this calculation is the correct one.

   (b) Explain how you would find the degrees of freedom in practice. What is the maximum possible degrees of freedom that you could have in this case? The minimum?

(c) Did the water conservation program have a statistically significant effect on water consumption at $\alpha = .05$? How do you know?

2. Researchers are studying whether there is ethnic favoritism by civil service employees in granting permits for new businesses. Out of 125 permit applications by members of Ethnic Group A, 68% were approved. Out of 88 permit applications by members of Ethnic Group B, 75% were approved.

   (a) Find the 95% confidence interval for the difference in the permit approval rates. Does it appear plausible that the observed difference is due to random chance? Explain.

   (b) If you were to perform a significance test for the difference in approval rates, what approval rate would represent the null hypothesis?

3. In another study of ethnic favoritism, researchers prepare a set of 20 permit applications for various types of businesses. In this set, the applicant names are associated with Ethnic Group A. They prepare second set of 20 applications that are identical except that the applicant names are associated with Ethnic Group B. For each set, they will measure the mean time that it takes for a decision to be made. What form of significance test should the researchers use?

4. A researcher is working with two variables. The first variable has 10 categories and the second variable that 3 categories. She constructs a cross-tabulation table. Then, she performs a $\chi^2$ test of independence and finds that the $\chi^2$ statistic is 26.01.

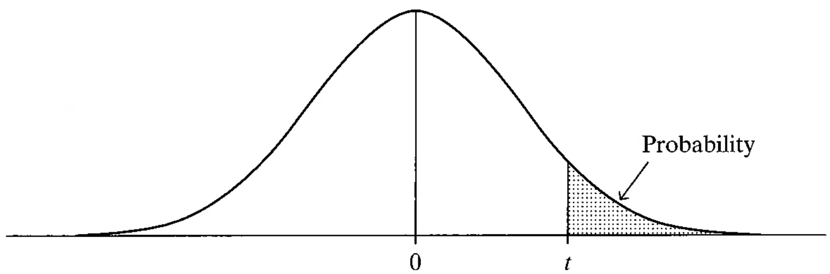
   (a) Can she reject the null hypothesis at the .05 level of significance? Explain.

   (b) Approximately what level of confidence does the researcher have in rejecting $H_0$?

5. A colleague asks you whether he should assume equal variances when running a t-test for the difference of means. What do you tell him and why?

6. When trying to identify whether there is a causal relationship between variables, what are the potential advantages of an experiment over an observational study?

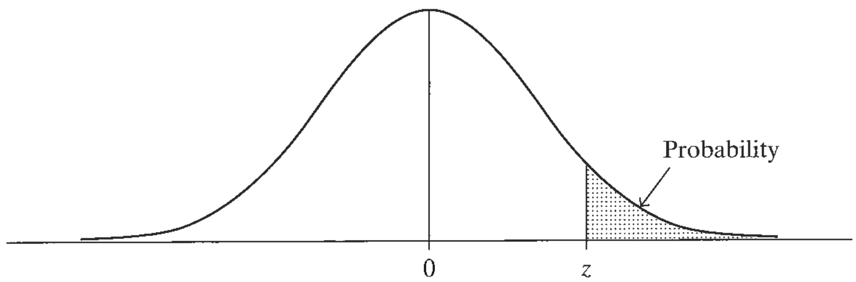
**TABLE B: *t* Distribution Critical Values**




| | Confidence Level | | | | | |
|---|---|---|---|---|---|---|
| | 80% | 90% | 95% | 98% | 99% | 99.8% |
| | Right-Tail Probability | | | | | |
| *df* | $t_{.100}$ | $t_{.050}$ | $t_{.025}$ | $t_{.010}$ | $t_{.005}$ | $t_{.001}$ |
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.656 | 318.289 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 22.328 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.214 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.894 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 | 3.686 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 | 3.646 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 | 3.611 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 | 3.579 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 | 3.552 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 | 3.527 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 | 3.505 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 | 3.485 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 | 3.467 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 | 3.450 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 | 3.435 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 | 3.421 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 | 3.408 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 | 3.396 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 | 3.385 |
| 40 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 | 3.307 |
| 50 | 1.299 | 1.676 | 2.009 | 2.403 | 2.678 | 3.261 |
| 60 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 | 3.232 |
| 80 | 1.292 | 1.664 | 1.990 | 2.374 | 2.639 | 3.195 |
| 100 | 1.290 | 1.660 | 1.984 | 2.364 | 2.626 | 3.174 |
| $\infty$ | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 | 3.091 |

Source: "Table of Percentage Points of the *t*-Distribution." Computed by Maxine Merrington, Biometrika, 32 (1941): 300. Reproduced by permission of the Biometrika trustees.

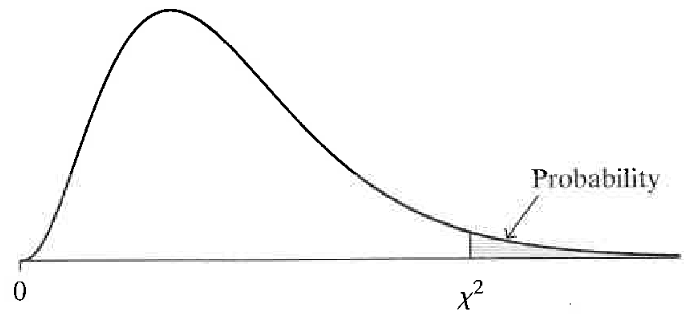
**TABLE A:** Normal curve tail probabilities. Standard normal probability in right-hand tail (for negative values of $z$, probabilities are found by symmetry)




| $z$ | Second Decimal Place of $z$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
| 0.0 | .5000 | .4960 | .4920 | .4880 | .4840 | .4801 | .4761 | .4721 | .4681 | .4641 |
| 0.1 | .4602 | .4562 | .4522 | .4483 | .4443 | .4404 | .4364 | .4325 | .4286 | .4247 |
| 0.2 | .4207 | .4168 | .4129 | .4090 | .4052 | .4013 | .3974 | .3936 | .3897 | .3859 |
| 0.3 | .3821 | .3783 | .3745 | .3707 | .3669 | .3632 | .3594 | .3557 | .3520 | .3483 |
| 0.4 | .3446 | .3409 | .3372 | .3336 | .3300 | .3264 | .3228 | .3192 | .3156 | .3121 |
| 0.5 | .3085 | .3050 | .3015 | .2981 | .2946 | .2912 | .2877 | .2843 | .2810 | .2776 |
| 0.6 | .2743 | .2709 | .2676 | .2643 | .2611 | .2578 | .2546 | .2514 | .2483 | .2451 |
| 0.7 | .2420 | .2389 | .2358 | .2327 | .2296 | .2266 | .2236 | .2206 | .2177 | .2148 |
| 0.8 | .2119 | .2090 | .2061 | .2033 | .2005 | .1977 | .1949 | .1922 | .1894 | .1867 |
| 0.9 | .1841 | .1814 | .1788 | .1762 | .1736 | .1711 | .1685 | .1660 | .1635 | .1611 |
| 1.0 | .1587 | .1562 | .1539 | .1515 | .1492 | .1469 | .1446 | .1423 | .1401 | .1379 |
| 1.1 | .1357 | .1335 | .1314 | .1292 | .1271 | .1251 | .1230 | .1210 | .1190 | .1170 |
| 1.2 | .1151 | .1131 | .1112 | .1093 | .1075 | .1056 | .1038 | .1020 | .1003 | .0985 |
| 1.3 | .0968 | .0951 | .0934 | .0918 | .0901 | .0885 | .0869 | .0853 | .0838 | .0823 |
| 1.4 | .0808 | .0793 | .0778 | .0764 | .0749 | .0735 | .0722 | .0708 | .0694 | .0681 |
| 1.5 | .0668 | .0655 | .0643 | .0630 | .0618 | .0606 | .0594 | .0582 | .0571 | .0559 |
| 1.6 | .0548 | .0537 | .0526 | .0516 | .0505 | .0495 | .0485 | .0475 | .0465 | .0455 |
| 1.7 | .0446 | .0436 | .0427 | .0418 | .0409 | .0401 | .0392 | .0384 | .0375 | .0367 |
| 1.8 | .0359 | .0352 | .0344 | .0336 | .0329 | .0322 | .0314 | .0307 | .0301 | .0294 |
| 1.9 | .0287 | .0281 | .0274 | .0268 | .0262 | .0256 | .0250 | .0244 | .0239 | .0233 |
| 2.0 | .0228 | .0222 | .0217 | .0212 | .0207 | .0202 | .0197 | .0192 | .0188 | .0183 |
| 2.1 | .0179 | .0174 | .0170 | .0166 | .0162 | .0158 | .0154 | .0150 | .0146 | .0143 |
| 2.2 | .0139 | .0136 | .0132 | .0129 | .0125 | .0122 | .0119 | .0116 | .0113 | .0110 |
| 2.3 | .0107 | .0104 | .0102 | .0099 | .0096 | .0094 | .0091 | .0089 | .0087 | .0084 |
| 2.4 | .0082 | .0080 | .0078 | .0075 | .0073 | .0071 | .0069 | .0068 | .0066 | .0064 |
| 2.5 | .0062 | .0060 | .0059 | .0057 | .0055 | .0054 | .0052 | .0051 | .0049 | .0048 |
| 2.6 | .0047 | .0045 | .0044 | .0043 | .0041 | .0040 | .0039 | .0038 | .0037 | .0036 |
| 2.7 | .0035 | .0034 | .0033 | .0032 | .0031 | .0030 | .0029 | .0028 | .0027 | .0026 |
| 2.8 | .0026 | .0025 | .0024 | .0023 | .0023 | .0022 | .0021 | .0021 | .0020 | .0019 |
| 2.9 | .0019 | .0018 | .0017 | .0017 | .0016 | .0016 | .0015 | .0015 | .0014 | .0014 |
| 3.0 | .00135 | | | | | | | | | |
| 3.5 | .000233 | | | | | | | | | |
| 4.0 | .0000317 | | | | | | | | | |
| 4.5 | .00000340 | | | | | | | | | |
| 5.0 | .000000287 | | | | | | | | | |

Source: R. E. Walpole, *Introduction to Statistics* (New York: Macmillan, 1968).

**TABLE C:** Chi-Squared Distribution Values for Various Right-Tail Probabilities




| *df* | Right-Tail Probability | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 | 0.001 |
| 1 | 1.32 | 2.71 | 3.84 | 5.02 | 6.63 | 7.88 | 10.83 |
| 2 | 2.77 | 4.61 | 5.99 | 7.38 | 9.21 | 10.60 | 13.82 |
| 3 | 4.11 | 6.25 | 7.81 | 9.35 | 11.34 | 12.84 | 16.27 |
| 4 | 5.39 | 7.78 | 9.49 | 11.14 | 13.28 | 14.86 | 18.47 |
| 5 | 6.63 | 9.24 | 11.07 | 12.83 | 15.09 | 16.75 | 20.52 |
| 6 | 7.84 | 10.64 | 12.59 | 14.45 | 16.81 | 18.55 | 22.46 |
| 7 | 9.04 | 12.02 | 14.07 | 16.01 | 18.48 | 20.28 | 24.32 |
| 8 | 10.22 | 13.36 | 15.51 | 17.53 | 20.09 | 21.96 | 26.12 |
| 9 | 11.39 | 14.68 | 16.92 | 19.02 | 21.67 | 23.59 | 27.88 |
| 10 | 12.55 | 15.99 | 18.31 | 20.48 | 23.21 | 25.19 | 29.59 |
| 11 | 13.70 | 17.28 | 19.68 | 21.92 | 24.72 | 26.76 | 31.26 |
| 12 | 14.85 | 18.55 | 21.03 | 23.34 | 26.22 | 28.30 | 32.91 |
| 13 | 15.98 | 19.81 | 22.36 | 24.74 | 27.69 | 29.82 | 34.53 |
| 14 | 17.12 | 21.06 | 23.68 | 26.12 | 29.14 | 31.32 | 36.12 |
| 15 | 18.25 | 22.31 | 25.00 | 27.49 | 30.58 | 32.80 | 37.70 |
| 16 | 19.37 | 23.54 | 26.30 | 28.85 | 32.00 | 34.27 | 39.25 |
| 17 | 20.49 | 24.77 | 27.59 | 30.19 | 33.41 | 35.72 | 40.79 |
| 18 | 21.60 | 25.99 | 28.87 | 31.53 | 34.81 | 37.16 | 42.31 |
| 19 | 22.72 | 27.20 | 30.14 | 32.85 | 36.19 | 38.58 | 43.82 |
| 20 | 23.83 | 28.41 | 31.41 | 34.17 | 37.57 | 40.00 | 45.32 |
| 25 | 29.34 | 34.38 | 37.65 | 40.65 | 44.31 | 46.93 | 52.62 |
| 30 | 34.80 | 40.26 | 43.77 | 46.98 | 50.89 | 53.67 | 59.70 |
| 40 | 45.62 | 51.80 | 55.76 | 59.34 | 63.69 | 66.77 | 73.40 |
| 50 | 56.33 | 63.17 | 67.50 | 71.42 | 76.15 | 79.49 | 86.66 |
| 60 | 66.98 | 74.40 | 79.08 | 83.30 | 88.38 | 91.95 | 99.61 |
| 70 | 77.58 | 85.53 | 90.53 | 95.02 | 100.4 | 104.2 | 112.3 |
| 80 | 88.13 | 96.58 | 101.8 | 106.6 | 112.3 | 116.3 | 124.8 |
| 90 | 98.65 | 107.6 | 113.1 | 118.1 | 124.1 | 128.3 | 137.2 |
| 100 | 109.1 | 118.5 | 124.3 | 129.6 | 135.8 | 140.2 | 149.5 |

Source: Calculated using *StaTable*, software from Cytel Software, Cambridge, MA.

# Public Policy 529 Formula Sheet

## Descriptive and Distributional Statistics

$$\bar{y} = \frac{\sum y_i}{n}$$

$$s^2 = \frac{\sum (y_i - \bar{y})^2}{n-1}$$

$$Z = \frac{y - \mu_y}{\sigma}$$

$$IQR = Q_3 - Q_1$$

$$SS = \sum (y_i - \bar{y})^2$$

$$s = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}$$

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}}$$

## Probability

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

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

$$P(\sim A) = 1 - P(A)$$

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

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$P(x) = \frac{n!}{x!(n-x)!} \pi^x (1-\pi)^{n-x}$$

## Confidence Intervals and Significance Tests

$$t = \frac{\bar{y} - \mu_0}{\hat{\sigma}_{\bar{y}}}$$

$$\hat{\sigma}_{\hat{\pi}} = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

$$Z = \frac{\hat{\pi} - \pi_0}{\hat{\sigma}_{\pi_0}}$$

$$\text{c.i.} = \bar{y} \pm t \cdot \hat{\sigma}_{\bar{y}}$$

$$\hat{\sigma}_{\bar{y}} = \frac{s}{\sqrt{n}}$$

$$\text{c.i.} = \bar{y} \pm Z \cdot \hat{\sigma}_{\bar{y}}$$

$$\hat{\sigma}_{\pi_0} = \sqrt{\frac{\pi_0(1-\pi_0)}{n}}$$

$$\text{c.i.} = \hat{\pi} \pm Z \cdot \hat{\sigma}_{\hat{\pi}}$$

$$se_{\text{diff}} = \sqrt{(se_1)^2 + (se_2)^2}$$

$$z = \frac{(\hat{\pi}_2 - \hat{\pi}_1) - H_0}{se_0}, \text{ where } se_0 = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n_1} + \frac{\hat{\pi}(1-\hat{\pi})}{n_2}} \qquad c.i. = (\hat{\pi}_2 - \hat{\pi}_1) \pm z\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}$$

---

$$t = \frac{(\bar{y}_2 - \bar{y}_1) - H_0}{se_{\text{diff}}} \qquad ci = (\bar{y}_2 - \bar{y}_1) \pm t \cdot se_{\text{diff}}$$

Unequal Variance:

$$se_{\text{diff}} = \hat{\sigma}_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} \qquad \text{approx. } df = \min(n_1 - 1, n_2 - 1)$$

Equal Variance:

$$se_{\text{diff}} = \hat{\sigma}_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{s_{pooled}^2}{n_1} + \frac{s_{pooled}^2}{n_2}} = s_{pooled}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$df = n_1 + n_2 - 2$$

---

$$t = \frac{\bar{y}_d - H_0}{\hat{\sigma}_{\bar{y}_d}} \qquad \hat{\sigma}_{\bar{y}_d} = \frac{s_d}{\sqrt{n}}, \text{ where } s_d = \sqrt{\frac{\sum(y_{di} - \bar{y}_d)^2}{n - 1}}$$

Measures of Association

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}, \quad \text{se}_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}, \text{ where } f_e = \frac{\text{(row total)(column total)}}{n}$$