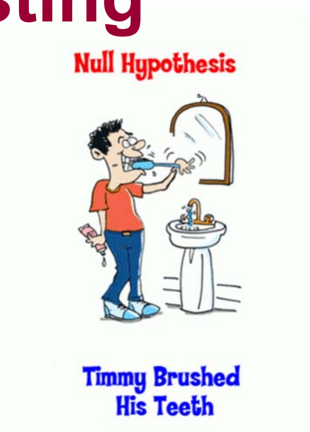
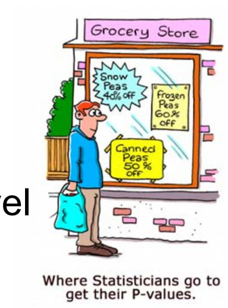
Hypothesis Testing



P-values



- The p-value is the smallest level of significance at which the H_0 would be rejected
  - $p(data|H_0)$ = p-value
- If p-value $\leq \alpha$, then reject H_0 at level α

  If p-value $> \alpha$, then do not reject H_0 at level α
- The lower the p-value, the stronger your evidence in support of alternative hypothesis

Example: p-Value

- The mean water volume is expected to be 20 Oz. Determine the mean water volume differs from 20 Oz assuming that the population STD to be 2 Oz
- A sample of size 36 finds the sample mean water volume to be 19 Oz
  - What is the p-value?
  - What is your conclusion?

Solution

- Step 1: Establish hypothesis
  - H_0: μ = 20 Oz
  - H_a: $\mu \neq$ 20 Oz
- Step 2: Determine appropriate statistical test and sampling distribution
  - a two-tailed test
  - σ is known: use z-distribution
- Step 3: Specify the Type I error rate (significance level)
  - α = 0.01

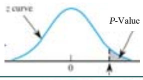
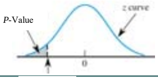
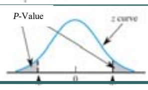
Solution, Cont'd

- Step 4: Gather data
  - n = 36, $\bar{X} = 19$
- Step 5: Calculate test statistic
  - $\mu_0 = 20$ Oz, $\sigma = 2$ Oz
  - $z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{19-20}{\frac{2}{\sqrt{36}}} = -3$
- Step 6: Calculate p-value
  - $\Phi(|z|) = \Phi(3)$
  - P-value = 2×(1-.9987) = 2×.0013 = .0026 < 0.01
- Step 7: State statistical conclusion
  - P-value < 0.01 Reject H_0

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	$\boxed{.9987}$	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

P-values

- H_0: $\mu = \mu_0$
- Our test statistic is: $z = \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$

Alternative Hypothesis		Rejection region for level α test
H_a: $\mu > \mu_0$		$P = p(Z>z)= 1-\Phi(z)$
H_a: $\mu < \mu_0$		$P = p(Z<z) = \Phi(z)$
H_a: $\mu \neq \mu_0$		$P = 2\ (1-\Phi(\lvert z \rvert))$

Example: p-Value

- The mean water volume is expected to be 20 Oz. Determine the mean water volume differs from 20 Oz
- A sample of size 36 finds the sample mean water volume to be 19 Oz and the sample STD to be 2 Oz
  - What is the p-value?
  - What is your conclusion?

Solution

- Step 1: Establish hypothesis
  - H_0: μ = 20 Oz
  - H_a: $\mu \neq$ 20 Oz
- Step 2: Determine appropriate statistical test and sampling distribution
  - a two-tailed test
  - σ is unknown, n<40: use t-distribution
- Step 3: Specify the Type I error rate (significance level)
  - α = 0.01

Solution, Cnt'd

- Step 4: Gather data
  - n = 36, $\bar{X}$ = 19
- Step 5: Calculate test statistic
  - μ_0 = 20 Oz, s = 2 Oz
  - $t = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}} = \frac{19-20}{\frac{2}{\sqrt{36}}} = -3$
- Step 6: Find the p-Value (See Table A.8)
  - t = -3.0, df=35
  - p-value = 2p(X>|-3|) = 2×0.002=0.004 < 0.01
- Step 7: State statistical conclusion
  - P-value < 0.01: Reject H_0

t \ ν	19	20	21	22	23	24	25	26	27	28	29	30	35	40	60	120	$\infty(=z)$
0.0	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500
0.1	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461	.461	.460	.460	.460	.460	.460
0.2	.422	.422	.422	.422	.422	.422	.422	.422	.421	.421	.421	.421	.421	.421	.421	.421	.421
0.3	.384	.384	.384	.383	.383	.383	.383	.383	.383	.383	.383	.383	.383	.383	.383	.382	.382
0.4	.347	.347	.347	.347	.346	.346	.346	.346	.346	.346	.346	.346	.346	.346	.345	.345	.345
0.5	.311	.311	.311	.311	.311	.311	.311	.311	.311	.310	.310	.310	.310	.310	.309	.309	.309
0.6	.278	.278	.278	.277	.277	.277	.277	.277	.277	.277	.277	.277	.276	.276	.275	.275	.274
0.7	.246	.246	.246	.246	.245	.245	.245	.245	.245	.245	.245	.245	.244	.244	.243	.243	.242
0.8	.217	.217	.216	.216	.216	.216	.216	.215	.215	.215	.215	.215	.215	.214	.213	.213	.212
0.9	.190	.189	.189	.189	.189	.189	.188	.188	.188	.188	.188	.188	.187	.187	.186	.185	.184
1.0	.165	.165	.164	.164	.164	.164	.163	.163	.163	.163	.163	.163	.162	.162	.161	.160	.159
1.1	.143	.142	.142	.142	.141	.141	.141	.141	.141	.140	.140	.140	.139	.139	.138	.137	.136
1.2	.122	.122	.122	.121	.121	.121	.121	.120	.120	.120	.120	.120	.119	.119	.117	.116	.115
1.3	.105	.104	.104	.104	.103	.103	.103	.103	.102	.102	.102	.102	.101	.101	.099	.098	.097
1.4	.089	.089	.088	.088	.087	.087	.087	.087	.086	.086	.086	.086	.085	.085	.083	.082	.081
1.5	.075	.075	.074	.074	.074	.073	.073	.073	.073	.072	.072	.072	.071	.071	.069	.068	.067
1.6	.063	.063	.062	.062	.062	.061	.061	.061	.061	.060	.060	.060	.059	.059	.057	.056	.055
1.7	.053	.052	.052	.052	.051	.051	.051	.051	.050	.050	.050	.050	.049	.048	.047	.046	.045
1.8	.044	.043	.043	.043	.042	.042	.042	.042	.042	.041	.041	.041	.040	.040	.038	.037	.036
1.9	.036	.036	.036	.035	.035	.035	.035	.034	.034	.034	.034	.034	.033	.032	.031	.030	.029
2.0	.030	.030	.029	.029	.029	.028	.028	.028	.028	.028	.027	.027	.027	.026	.025	.024	.023
2.1	.025	.024	.024	.024	.023	.023	.023	.023	.023	.022	.022	.022	.022	.021	.020	.019	.018
2.2	.020	.020	.020	.019	.019	.019	.019	.018	.018	.018	.018	.018	.017	.017	.016	.015	.014
2.3	.016	.016	.016	.016	.015	.015	.015	.015	.015	.015	.014	.014	.014	.013	.012	.012	.011
2.4	.013	.013	.013	.013	.012	.012	.012	.012	.012	.012	.012	.011	.011	.011	.010	.009	.008
2.5	.011	.011	.010	.010	.010	.010	.010	.010	.009	.009	.009	.009	.009	.008	.008	.007	.006
2.6	.009	.009	.008	.008	.008	.008	.008	.008	.007	.007	.007	.007	.007	.007	.006	.005	.005
2.7	.007	.007	.007	.007	.006	.006	.006	.006	.006	.006	.006	.006	.005	.005	.004	.004	.003
2.8	.006	.006	.005	.005	.005	.005	.005	.005	.005	.005	.005	.004	.004	.004	.003	.003	.003
2.9	.005	.004	.004	.004	.004	.004	.004	.004	.004	.004	.004	.003	.003	.003	.003	.002	.002
3.0	.004	.004	.003	.003	.003	.003	.003	.003	.003	.003	.003	.003	$\boxed{.002}$	.002	.002	.002	.001
3.1	.003	.003	.003	.003	.003	.002	.002	.002	.002	.002	.002	.002	.002	.002	.001	.001	.001
3.2	.002	.002	.002	.002	.002	.002	.002	.002	.002	.002	.002	.002	.001	.001	.001	.001	.001
3.3	.002	.002	.002	.002	.002	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.000
3.4	.002	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.000	.000
3.5	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.000	.000	.000
3.6	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.001	.000	.000	.000	.000	.000
3.7	.001	.001	.001	.001	.001	.001	.001	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000
3.8	.001	.001	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
3.9	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
4.0	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000

P-value

- H_0: $\mu = \mu_0$
- Our test statistic is: $t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$

Alternative Hypothesis	Rejection region for level α test
H_a: $\mu > \mu_0$	P = p(X>t)
H_a: $\mu < \mu_0$	P = p(X<t)
H_a: $\mu \neq \mu_0$	P = 2p(X>|t|)

Relationship between CI and Hypothesis Test

- The mean water volume is expected to be 20 Oz. Determine the mean water volume differs from 20 Oz
- A sample of size 36 finds the sample mean water volume to be 19 Oz and the sample STD to be 2 Oz
  - What is 99% two-sided CI?
  - Compare the result with the p-Value

Relationship between CI and Hypothesis Test

- 99% two-sided CI for the water volume

$$\bar{x} \pm t_{\frac{\alpha}{2},n-1}\frac{s}{\sqrt{n}} = \bar{x} \pm t_{0.005,36-1}\frac{s}{\sqrt{n}}$$

$$= 19 \pm 2.724\frac{2}{\sqrt{36}}$$

$$= 19 \pm 0.908$$

- μ_0 = 20 is not contained within this CI, which is consistent with the hypothesis testing problem having a p-value of 0.004, so that the null hypothesis is rejected at size α=0.01

df/α =	.40	.25	.10	.05	.025	.01	.005	.001	.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646
35	0.255	0.682	1.306	1.690	2.030	2.438	2.724	3.340	3.591
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	0.255	0.679	1.299	1.676	2.009	2.403	2.678	3.261	3.496
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	3.160	3.373
inf.	0.253	0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Example
Hypothesis Testing

- The mean length of a part is expected to be 30mm
- Determine the mean length differs from 30 mm assuming that the population STD of 2mm
- A sample of size 36 finds the sample mean length to be 29 mm
- Is this difference statistically significant at a significance level of .05?

Solution

- Step 1: Establish hypothesis
  - H_0: μ = 30mm
  - H_a: $\mu \neq$ 30mm
- Step 2: Determine appropriate statistical test and sampling distribution
  - a two-tailed test
  - σ is known: use z-distribution
- Step 3: Specify the Type I error rate (significance level)
  - α = 0.05

Solution, Cnt'd

- Step 4: State the decision rule
  - If z > = $z_{.025}$), reject H_0
  - If z < = $-z_{.025}$), reject H_0
- Step 5: Gather data
  - n = 36, $\bar{X}$ = 29
- Step 6: Calculate test statistic
  - μ_0 = 30mm, σ = 2mm
  - $z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{29-30}{\frac{2}{\sqrt{36}}} = -3$
- Step 7: State statistical conclusion
  - Z = -3 < -1.96: reject the H_0 at the 5% level
  - It is very unlikely that the mean is actually 30 mm

Table 4.1 Standard Normal Percentiles and Critical Values

Percentile	90	95	97.5	99	99.5	99.9	99.95
α (tail area)	.1	.05	.025	.01	.005	.001	.0005
$z_\alpha = 100(1-\alpha)$th percentile	1.28	1.645	1.96	2.33	2.58	3.08	3.27

Example: p-Value

- The mean length of a part is expected to be 30mm. Determine the mean length differs from 30 mm assuming that the population STD of 2mm
- A sample of size 36 finds the sample mean length to be 29 mm
  - What is the p-value?

Solution

- Step 1: Establish hypothesis
  - H_0: μ = 30mm
  - H_a: $\mu \neq$ 30mm
- Step 2: Determine appropriate statistical test and sampling distribution
  - a two-tailed test
  - σ is known: use z-distribution
- Step 3: Specify the Type I error rate (significance level)
  - α = 0.05

Solution, Cnt'd

- Step 4: Gather data
  - n = 36, $\bar{X}$ = 29
- Step 5: Calculate test statistic
  - μ_0 = 30mm, σ = 2mm
  - $z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{29-30}{\frac{2}{\sqrt{36}}} = -3$
- Step 6: Calculate p-value
  - $\Phi(|z|) = \Phi(3)$ = .9987
  - P-value = 2×(1-.9987) = 2×.0013 = .0026
- Step 7: State statistical conclusion
  - P-value < 0.05: Reject H_0