Ch. 7. “One sample” hypothesis tests for μ and σ

Prof. Tesler

Math 186
Winter 2019

Introduction

- Consider the SAT math scores again.
  Secretly, the mean is 500 and the standard deviation is 100.
- **Chapter 5:** We assumed $\sigma = 100$ was known. We estimated μ from data as a confidence interval centered on the sample mean.
- **Chapter 6:** We did hypothesis tests about μ under the same circumstances.
- **Chapter 7:** Both μ and σ are unknown. We estimate both of them from data, either for confidence intervals or hypothesis tests.

Data

Exp. #	Values $x_1, \ldots, x_6$	Sample mean m	Sample Var. s^2	Sample SD s
#1	650, 510, 470, 570, 410, 370	496.67	10666.67	103.28
#2	510, 420, 520, 360, 470, 530	468.33	4456.67	66.76
#3	470, 380, 480, 320, 430, 490	428.33	4456.67	66.76

Number of standard deviations m is away from μ when $\mu = 500$ and $\sigma = 100$, for sample mean of $n = 6$ points

Number of standard deviations if σ is known:

The z-score of m is

$$z = \frac{m - \mu}{\sigma / \sqrt{n}} = \frac{m - 500}{100 / \sqrt{6}}$$

Estimating number of standard deviations if σ is unknown:

The t-*score* of m is

$$t = \frac{m - \mu}{s / \sqrt{n}} = \frac{m - 500}{s / \sqrt{6}}$$

- It uses sample standard deviation s in place of σ.
- s is computed from the same data as m. So for t, numerator and denominator depend on data, while for z, only numerator does.
- t has the same degrees of freedom as s; here, $df = n - 1 = 5$.
- The random variable is called T_5 (T distribution with 5 degrees of freedom).

Number of standard deviations m is away from μ

Data

Exp. #	Values $x_1, \ldots, x_6$	Sample mean m	Sample Var. s^2	Sample SD s
#1	650, 510, 470, 570, 410, 370	496.67	10666.67	103.28
#2	510, 420, 520, 360, 470, 530	468.33	4456.67	66.76
#3	470, 380, 480, 320, 430, 490	428.33	4456.67	66.76

#1: $z = \dfrac{496.67 - 500}{100/\sqrt{6}} \approx -.082 \qquad t = \dfrac{496.67 - 500}{103.28/\sqrt{6}} \approx -.079$ **Close**

#2: $z = \dfrac{468.33 - 500}{100/\sqrt{6}} \approx -.776 \qquad t = \dfrac{468.33 - 500}{66.76/\sqrt{6}} \approx -1.162$ **Far**

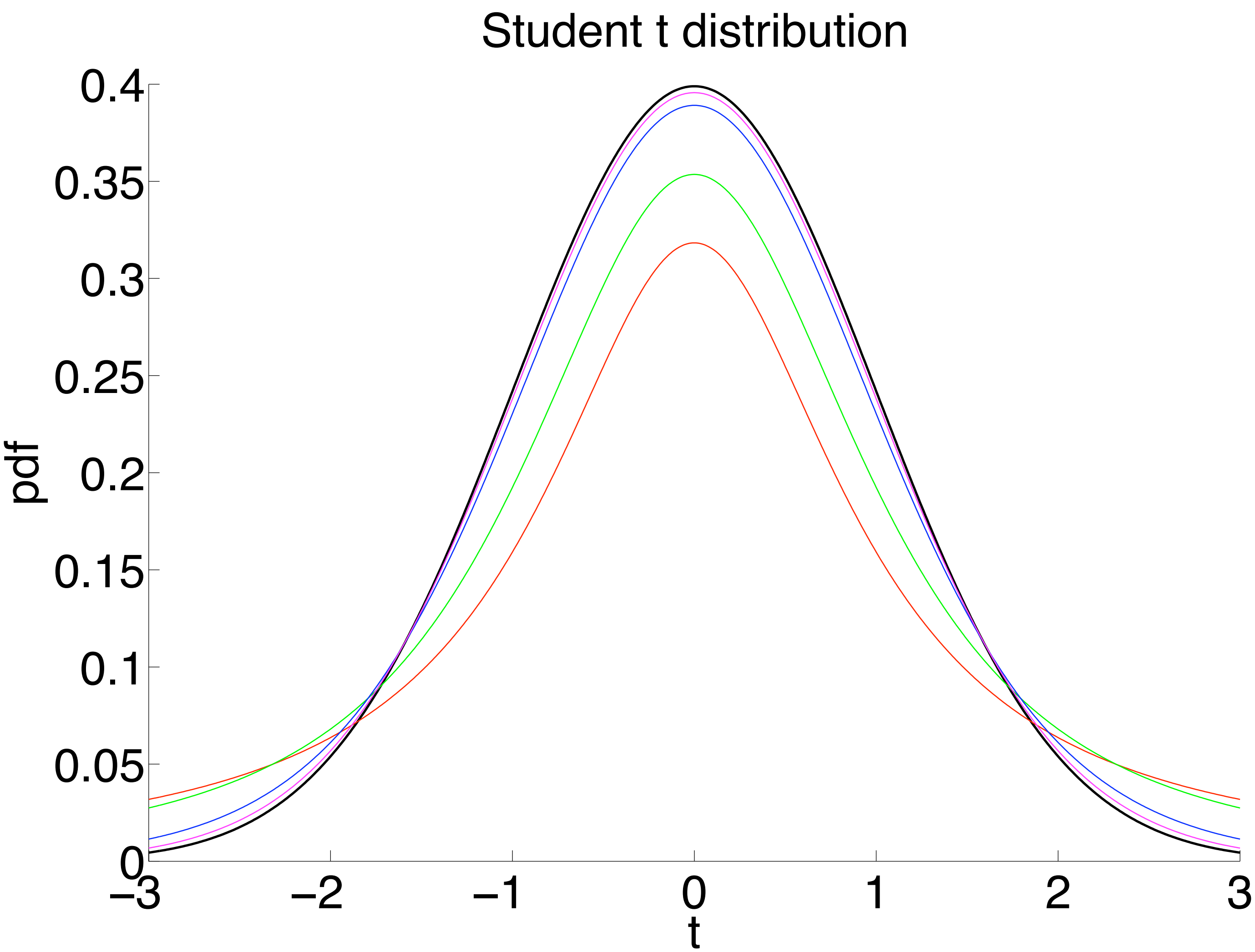
#3: $z = \dfrac{428.33 - 500}{100/\sqrt{6}} \approx -1.756 \qquad t = \dfrac{428.33 - 500}{66.76/\sqrt{6}} \approx -2.630$ **Far**

Student t distribution

- In $z = \frac{m-\mu}{\sigma/\sqrt{n}}$, the numerator depends on $x_1, \ldots, x_n$ while the denominator is constant.
  But in $t = \frac{m-\mu}{s/\sqrt{n}}$, both the numerator and denominator are functions of $x_1, \ldots, x_n$ (since m and s are functions of them).
- The pdf of t is no longer the standard normal distribution, but instead is a new distribution, T_{n-1}, the t-distribution with $n-1$ degrees of freedom. $(d.f. = n-1)$
- The pdf is still symmetric and "bell-shaped," *but not the same "bell" as the normal distribution.*
- Degrees of freedom $d.f.=n-1$ match here and in the s^2 formula.
- As $d.f.$ rises, the curves get closer to the standard normal curve; the curves are really close for $d.f. \geqslant 30$.
- This was developed in 1908 by William Gosset under the pseudonym "Student." He worked at Guinness Brewery with small sample sizes, such as $n = 3$.

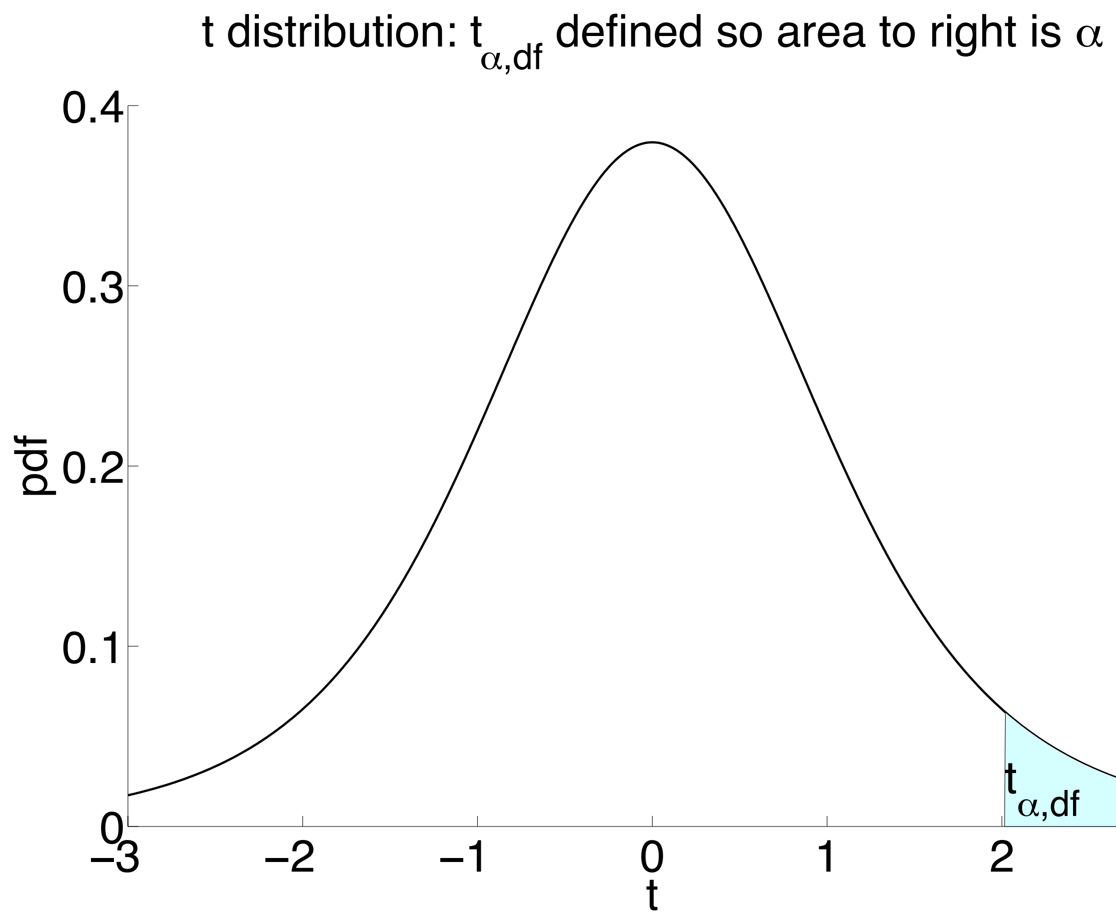
Student t distribution

The curves from bottom to top (at $t = 0$) are for $d.f. = 1, 2, 10, 30$.
The top one is the standard normal curve:



Student t distribution

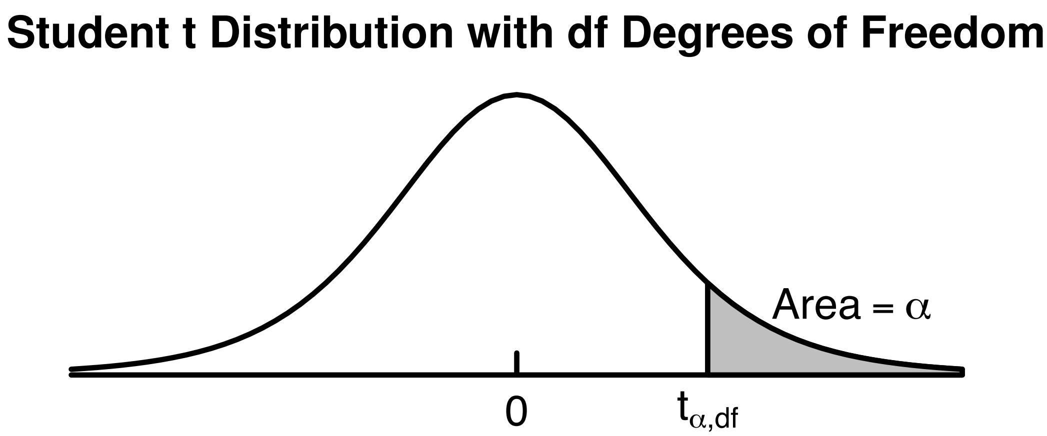
- For the t-distribution with df degrees of freedom (random variable T_{df}), define $t_{\alpha,df}$ so that $P(T_{df} \geqslant t_{\alpha,df}) = \alpha$.



- This is analogous to the standard normal distribution, where z_α was defined so the area right of z_α is α: $P(Z \geqslant z_\alpha) = \alpha$.

See t table in the back of the book (Table A.2)

Look up $t_{.025,5} = 2.5706$



α

df	0.20	0.15	0.10	0.05	0.025	0.01	0.005
1	1.3764	1.9626	3.0777	6.3138	12.7062	31.8205	63.6567
2	1.0607	1.3862	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.9785	1.2498	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.9410	1.1896	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.9195	1.1558	1.4759	2.0150	**2.5706**	3.3649	4.0321
6	0.9057	1.1342	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.8960	1.1192	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.8889	1.1081	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.8834	1.0997	1.3830	1.8331	2.2622	2.8214	3.2498

Note: Rounding and # decimals is different than the book's version.

Confidence intervals for estimating μ from m

- In Chapter 5, we made 95% confidence intervals for μ from m assuming we knew σ (and it works for any n):

$$\left(m - 1.96\frac{\sigma}{\sqrt{n}}, m + 1.96\frac{\sigma}{\sqrt{n}}\right)$$

- We now replace σ by the estimate s from the data.
  1.96 is replaced by a cutoff for t for $6 - 1 = 5$ degrees of freedom.

  To put 95% of the area in the center, 2.5% on the left, and 2.5% on the right, look up $t_{.025,5} = 2.5706$ in the table in the book.

$$\left(m - \frac{2.5706\,s}{\sqrt{6}}, m + \frac{2.5706\,s}{\sqrt{6}}\right)$$

- Note that the cutoff 2.5706 depended on $df = n - 1 = 5$ and would change for other n's; also, we still divide by $\sqrt{n} = \sqrt{6}$.

Confidence intervals for estimating μ from m

Formulas for 2-sided $100(1-\alpha)\%$ confidence interval for μ

When σ is known, use normal distribution

$$\left(m - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, m + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}\right)$$

95% confidence interval ($\alpha = 0.05$) with $\sigma = 100$, $z_{.025} = 1.96$:

$$\left(m - \frac{1.96(100)}{\sqrt{n}}, m + \frac{1.96(100)}{\sqrt{n}}\right)$$

When σ is not known, and m, s estimated from same n points

$$\left(m - \frac{t_{\alpha/2,n-1} \cdot s}{\sqrt{n}}, m + \frac{t_{\alpha/2,n-1} \cdot s}{\sqrt{n}}\right)$$

A 95% confidence interval ($\alpha = .05$) when $n = 6$; $t_{.025,5} = 2.5706$

$$\left(m - \frac{2.5706\, s}{\sqrt{6}}, m + \frac{2.5706\, s}{\sqrt{6}}\right)$$

- The cutoff $z = 1.96$ doesn't depend on n, but $t = 2.5706$ does ($df = n - 1 = 5$) and would change for other values of n.
- In both versions, we divide by $\sqrt{n} = \sqrt{6}$.

95% confidence intervals for μ

Exp. #	Data $x_1, \ldots, x_6$	m	s^2	s
#1	650, 510, 470, 570, 410, 370	496.67	10666.67	103.28
#2	510, 420, 520, 360, 470, 530	468.33	4456.67	66.76
#3	470, 380, 480, 320, 430, 490	428.33	4456.67	66.76

When σ known (say $\sigma = 100$), use normal distribution

#1: $(496.67 - \frac{1.96(100)}{\sqrt{6}}, 496.67 + \frac{1.96(100)}{\sqrt{6}}) = (416.65, 576.69)$

#2: $(468.33 - \frac{1.96(100)}{\sqrt{6}}, 468.33 + \frac{1.96(100)}{\sqrt{6}}) = (388.31, 548.35)$

#3: $(428.33 - \frac{1.96(100)}{\sqrt{6}}, 428.33 + \frac{1.96(100)}{\sqrt{6}}) = (348.31, 508.35)$

When σ not known, estimate σ by s and use t-distribution

#1: $(496.67 - \frac{2.5706(103.28)}{\sqrt{6}}, 496.67 + \frac{2.5706(103.28)}{\sqrt{6}}) = (388.28, 605.06)$

#2: $(468.33 - \frac{2.5706(66.76)}{\sqrt{6}}, 468.33 + \frac{2.5706(66.76)}{\sqrt{6}}) = (398.27, 538.39)$

#3: $(428.33 - \frac{2.5706(66.76)}{\sqrt{6}}, 428.33 + \frac{2.5706(66.76)}{\sqrt{6}}) = (358.27, 498.39)$
(missing 500)

Hypothesis tests for μ

Test H_0: $\mu = 500$ vs. H_1: $\mu \neq 500$ at significance level $\alpha = .05$

Exp. #	Data $x_1, \ldots, x_6$	m	s^2	s
#1	650, 510, 470, 570, 410, 370	496.67	10666.67	103.28
#2	510, 420, 520, 360, 470, 530	468.33	4456.67	66.76
#3	470, 380, 480, 320, 430, 490	428.33	4456.67	66.76

When σ is known (say $\sigma = 100$)

Reject H_0 when $|z| \geqslant z_{\alpha/2} = z_{.025} = 1.96$.

#1: $z = -.082$, $|z| < 1.96$ so accept H_0.
#2: $z = -.776$, $|z| < 1.96$ so accept H_0.
#3: $z = -1.756$, $|z| < 1.96$ so accept H_0.

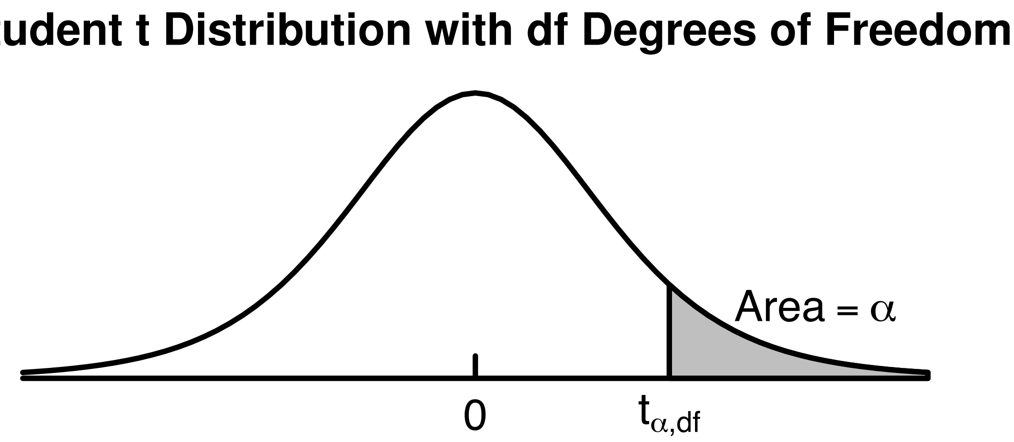
When σ is not known, but is estimated by s

Reject H_0 when $|t| \geqslant t_{\alpha/2,n-1} = t_{.025,5} = 2.5706$.

#1: $t = -.079$, $|t| < 2.5706$ so accept H_0.
#2: $t = -1.162$, $|t| < 2.5706$ so accept H_0.
#3: $t = -2.630$, $|t| \geqslant 2.5706$ so reject H_0.

See t table in the back of the book (Table A.2)

Use to compute approx. 2-sided P-values for $t = -.079, -1.162, -2.630, \ d.f. = 5$



α

df	0.20	0.15	0.10	0.05	0.025	0.01	0.005
1	1.3764	1.9626	3.0777	6.3138	12.7062	31.8205	63.6567
2	1.0607	1.3862	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.9785	1.2498	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.9410	1.1896	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.9195	1.1558	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.9057	1.1342	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.8960	1.1192	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.8889	1.1081	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.8834	1.0997	1.3830	1.8331	2.2622	2.8214	3.2498

t-test using P-values

Exp. #	Data $x_1, \ldots, x_6$	m	s^2	s
#1	650, 510, 470, 570, 410, 370	496.67	10666.67	103.28
#2	510, 420, 520, 360, 470, 530	468.33	4456.67	66.76
#3	470, 380, 480, 320, 430, 490	428.33	4456.67	66.76

Test H_0: $\mu = 500$ vs. H_1: $\mu \neq 500$ at significance level $\alpha = .05$.

- #1: $P = P(T \leqslant -.079) + P(T \geqslant .079) > 2(.20) = .40$
  so $P > \alpha$ ($P > .40 > .05$) and we accept H_0.
- #2: $P = P(T \leqslant -1.16) + P(T \geqslant 1.16) \approx 2(.15) = .30$.
  Since $P > \alpha$ ($.30 > .05$), accept H_0.
- #3: $P = P(T \leqslant -2.63) + P(T \geqslant 2.63)$

  P is between $2(.025) = .05$ and $2(.01) = .02$ based on the table.
  So $P \leqslant .05$ and we reject H_0.
- **On a calculator:** #1: $P = .9401$ #2: $P = .2977$ #3: $P = .0465$

7.5. The χ^2 ("Chi-squared") distribution

Hypothesis tests for σ^2

The χ^2 ("Chi-squared") distribution (Chapter 7.5)

- We'll do a hypothesis test for the variance, σ^2, of the normal distribution, just like we did for the mean, μ:

$$H_0: \sigma^2 = {\sigma_0}^2 \qquad \text{vs.} \qquad H_1: \sigma^2 \neq {\sigma_0}^2$$

Example: $$H_0: \sigma^2 = 10000 \qquad \text{vs.} \qquad H_1: \sigma^2 \neq 10000$$

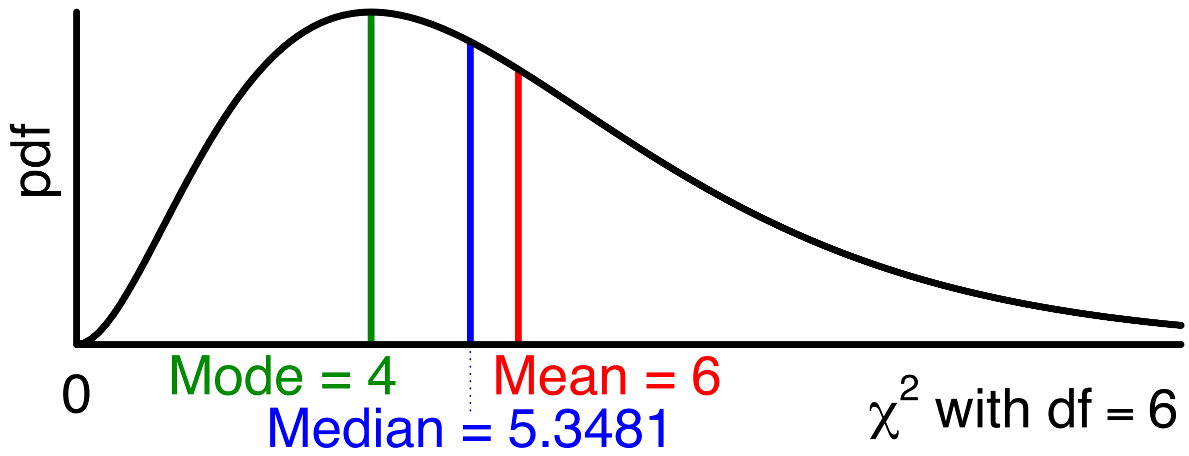
- Sample variance s^2 estimates theoretical variance σ^2.
  Use the ratio $s^2/{\sigma_0}^2$ to test consistency with H_0.
  Given a sample of size n, compute s^2, and plug it into this formula:

$$\text{Chi-squared: } \chi^2 = \frac{(n-1)s^2}{{\sigma_0}^2} = \sum_{i=1}^{n} \frac{(x_i - m)^2}{{\sigma_0}^2}$$

Degrees of freedom: $df = n - 1$ (same as for t)

- This test statistic is called *Chi-squared*.
  Note that χ and x are different.
  χ is the Greek letter chi. The data is x_i, with the letter x.

The χ^2 ("Chi-squared") distribution (Chapter 7.5)



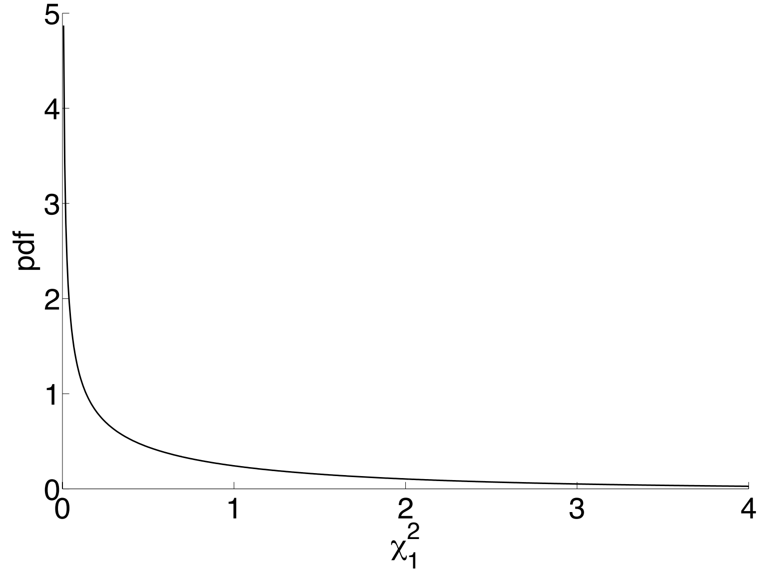
- The chi-squared distribution with k degrees of freedom has

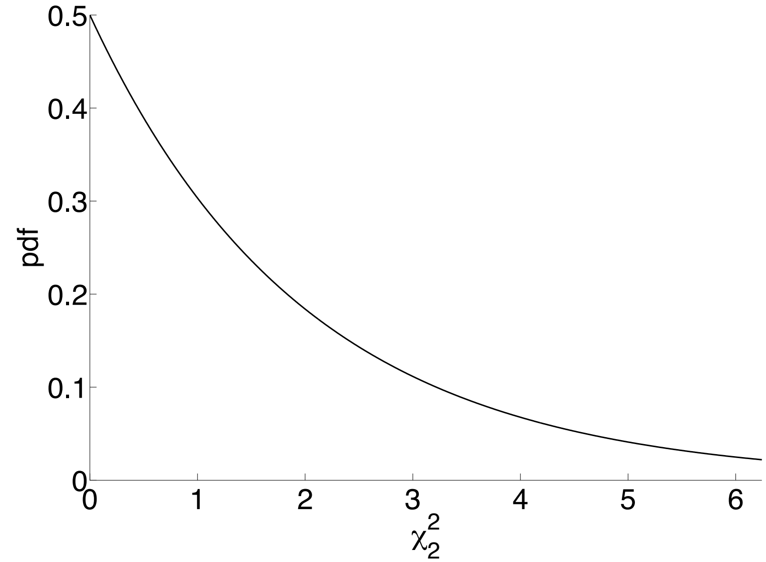
Range	$[0, \infty)$
Mean	$\mu = k$
Variance	$\sigma^2 = 2k$
Mode	$\chi^2 = k - 2$ (the pdf is maximum for $\chi^2 = k - 2$)
Median	Between k and $k - \frac{2}{3}$. Asymptotically decreases $\to k - \frac{2}{3}$ as $k \to \infty$.

- Unlike z and t, the pdf for χ^2 is NOT symmetric, and the mean, median, and mode are different.

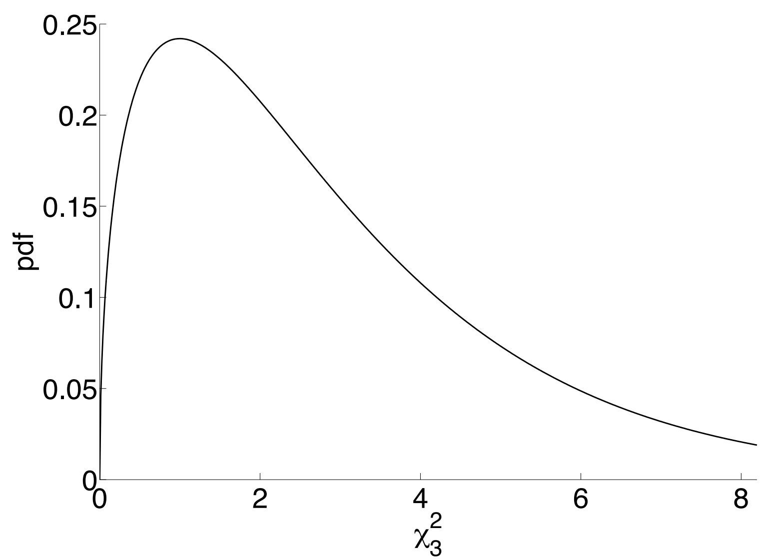
χ^2 ("Chi-squared") distribution — pdf graphs

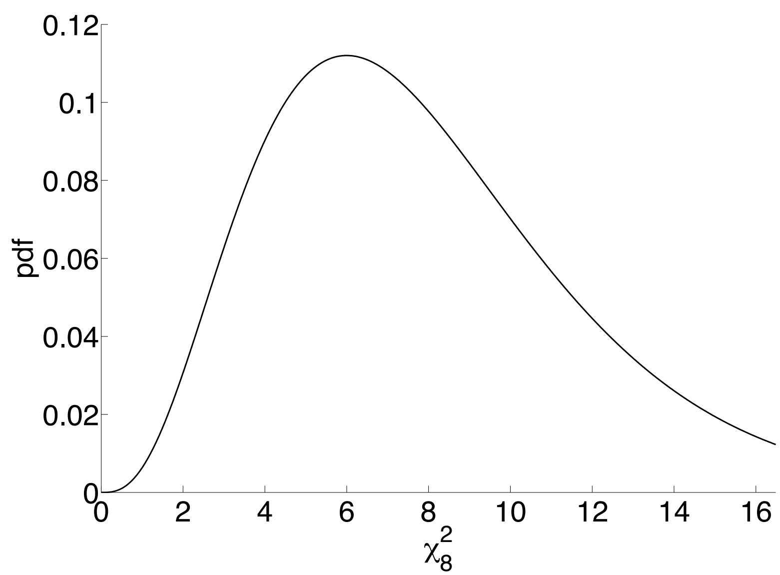
The graphs for 1 and 2 degrees of freedom are decreasing:



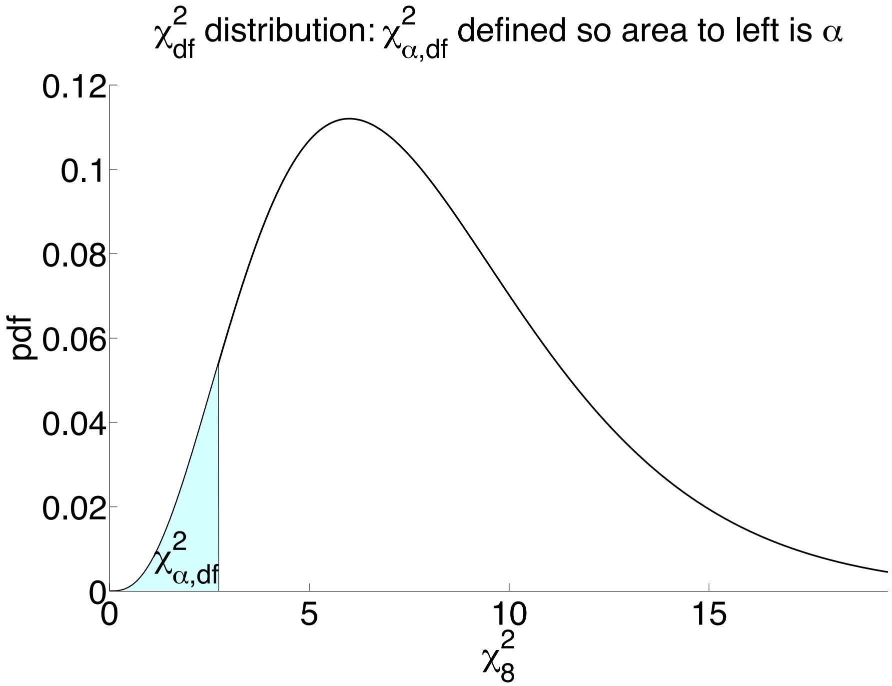


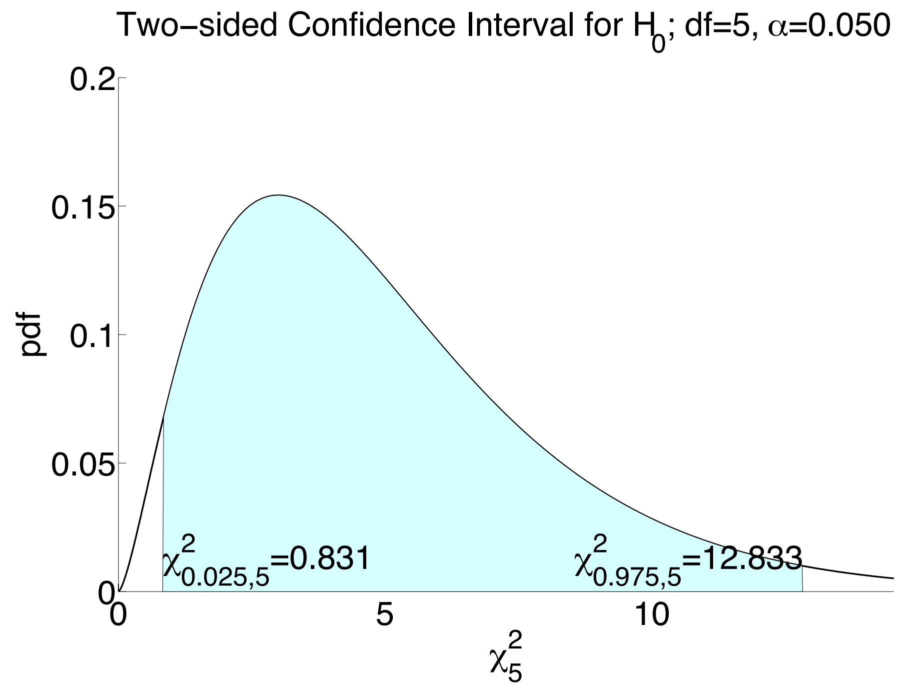
The rest are "hump" shaped and skewed to the right:





χ^2 ("Chi-squared") distribution — Cutoffs





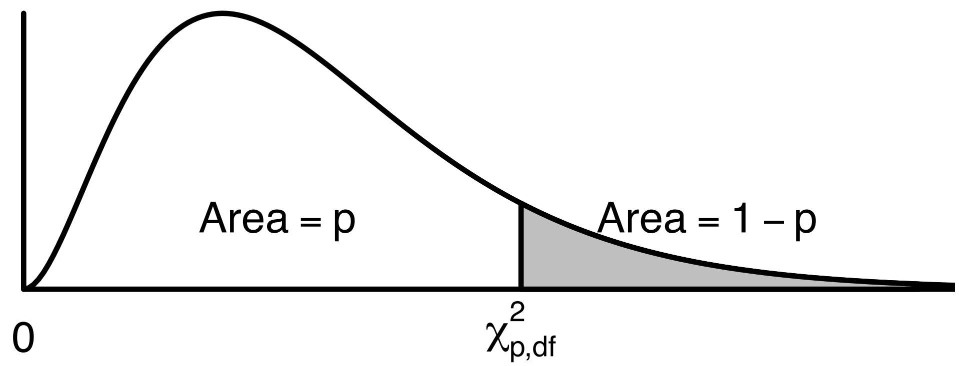
- Define $\chi^2_{\alpha,df}$ as the number where the cdf (area *left* of it) is α:
$$P(\chi^2_{df} \leqslant \chi^2_{\alpha,df}) = \alpha$$
- This is different than how our book did it for the z and t-distributions, because this pdf isn't symmetric.
- We still put 95% of the area in the middle and 2.5% at each end, but the lower and upper cutoffs are determined separately instead of $\pm$ each other.

See χ^2 table in the back of the book (Table A.3)

For two-sided test with $\alpha = .05$ and $n = 6$, look up

$$\chi^2_{\alpha/2,n-1} = \chi^2_{.025,5} = .831 \quad \text{and} \quad \chi^2_{1-\alpha/2,n-1} = \chi^2_{.975,5} = 12.832$$

χ^2 Distribution with df Degrees of Freedom



p

df	0.010	0.025	0.050	0.10	0.90	0.95	0.975	0.99
1	0.000157	0.000982	0.00393	0.015	2.705	3.841	5.023	6.634
2	0.020	0.050	0.102	0.210	4.605	5.991	7.377	9.210
3	0.114	0.215	0.351	0.584	6.251	7.814	9.348	11.344
4	0.297	0.484	0.710	1.063	7.779	9.487	11.143	13.276
5	0.554	**0.831**	1.145	1.610	9.236	11.070	**12.832**	15.086
6	0.872	1.237	1.635	2.204	10.644	12.591	14.449	16.811
7	1.239	1.689	2.167	2.833	12.017	14.067	16.012	18.475
8	1.646	2.179	2.732	3.489	13.361	15.507	17.534	20.090
9	2.087	2.700	3.325	4.168	14.683	16.918	19.022	21.665

Two-sided hypothesis test for variance

Test $H_0 : \sigma^2 = {\sigma_0}^2$ vs. $H_1 : \sigma^2 \neq {\sigma_0}^2$

Decision procedure

Test $H_0 : \sigma^2 = 10000$ vs. $H_1 : \sigma^2 \neq 10000$ at sig. level $\alpha = .05$ (so $\sigma_0 = 100$)

1. Get a sample $x_1, \ldots, x_n$.

   650, 510, 470, 570, 410, 370 with $n = 6$

2. Calculate $m = \frac{x_1 + \cdots + x_n}{n}$ and $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2$.

   $m = 496.67, \quad s^2 = 10666.67, \quad s = 103.28$

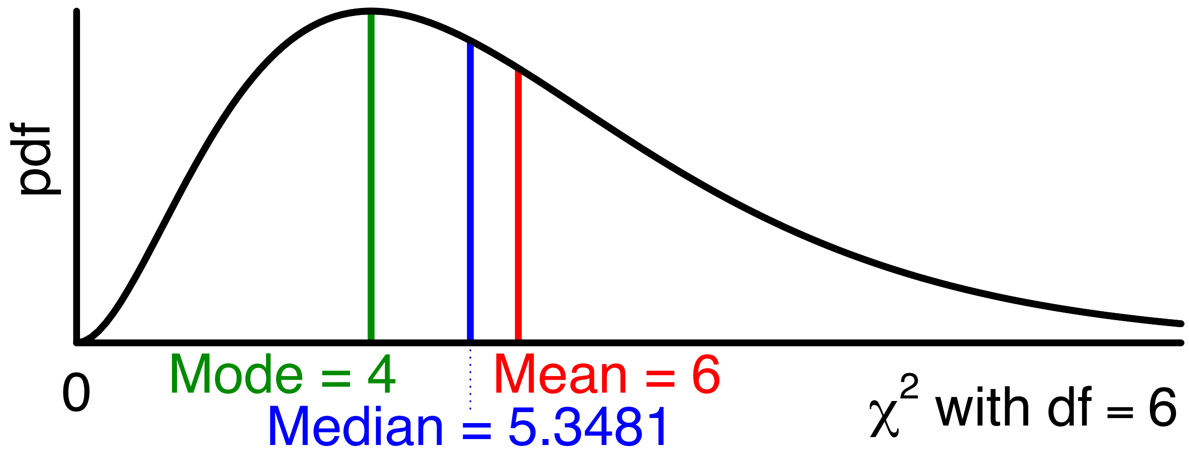
3. Calculate the test-statistic $\chi^2 = \frac{(n-1)s^2}{{\sigma_0}^2} = \sum_{i=1}^{n} \frac{(x_i - m)^2}{{\sigma_0}^2}$

   $\chi^2 = \frac{(6-1)(10666.67)}{10000} = 5.33$

4. Accept H_0 if χ^2 is between $\chi^2_{\alpha/2,n-1}$ and $\chi^2_{1-\alpha/2,n-1}$.
   Reject H_0 otherwise.

   $\chi^2_{.025,5} = .831$ and $\chi^2_{.975,5} = 12.832$.
   Since $\chi^2 = 5.33$ is between these, we accept H_0.
   (Or, there is insufficient evidence to reject $\sigma^2 = 10000$.)

Mean, Median, and Mode of χ^2



- $H_0 : \sigma^2 = {\sigma_0}^2$ vs. $H_1 : \sigma^2 \neq {\sigma_0}^2$
- Unlike z and t, the mean, median, and mode of χ^2 are different.

  Mean $\mu = k$ Median $\approx k - 2/3$ Mode $\chi^2 = k - 2$

  Question: Which of these should χ^2 be close to if H_0 holds?
- **Answer:** The median.
  - The hypothesis test cutoffs and P-values are based on the cdf.
  - The median is based on the cdf (it's where the cdf equals 1/2), while mean and mode are not.
  - The median is regarded as most consistent with H_0.

Properties of Chi-squared distribution

1. **Definition of Chi-squared distribution:**

   Let $Z_1, \ldots, Z_k$ be independent standard normal variables.

   Let $\chi_k^2 = {Z_1}^2 + \cdots + {Z_k}^2$.

   The pdf of the random variable χ_k^2 is the "chi-squared distribution with k degrees of freedom."

   The book has the exact formula of the pdf (but you don't need to know it).

2. **Pooling property:** If X and Y are independent χ^2 random variables with k and m degrees of freedom respectively, then $X+Y$ is a χ^2 random variable with $k+m$ degrees of freedom.