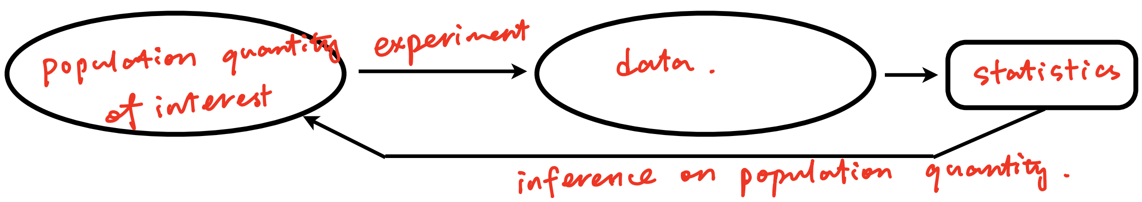
Sampling Distributions and the Central Limit Theorem **Notes ~~10~~ 9**

Associated Reading: Wackerly 7, Chapter 7, Sections 1-4

This chapter will conclude the discussion of functions of random variables that began in Chapter 5, and lay the last groundwork that you need before learning about estimators, confidence intervals, and hypothesis testing in Chapters 8-10.

The meta-idea here is that you've sampled iid r.v.'s $\{Y_1, Y_2, \cdots, Y_n\}$ from some population with unknown parameters, parameters that you'd like to estimate by examining functions of the r.v.'s. Here we remind ourselves of a definition:

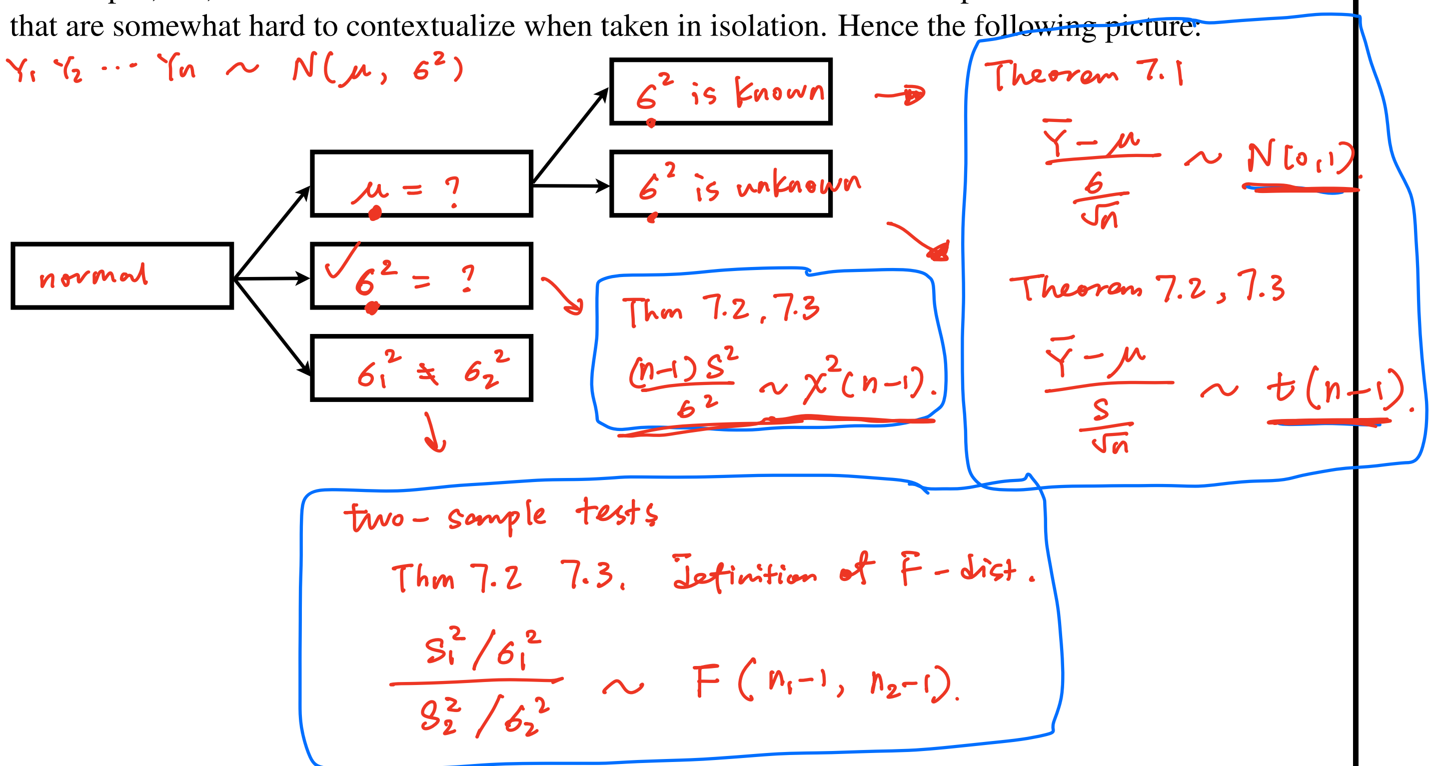
statistic : a funtion of data. i.e. function of r.v.s.



Just as a random variable is drawn from a pmf/pdf, *a statistic is drawn from a* **sampling** *distribution*, which is derivable from the pmf's or pdf's of the individual data. Defining the concept of the sampling distribution, and indicating how one may estimate it via simulation, is the subject of Section 7.1.

$\propto$

In Section 7.2, we look specifically at cases where the sampling distributions are derivable from the normal pdf, i.e., all our data are iid normal r.v.'s. There are three theorems presented in this section that are somewhat hard to contextualize when taken in isolation. Hence the following picture:



$\boxed{1^\circ}$ sampling distribution: $\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$.

We have seen Theorem 7.1 previously:

- If $\bar{Y} = (1/n)\sum_{i=1}^{n} Y_i$, where the Y_i's are iid samples from $N(\mu, \sigma^2)$, then $\bar{Y} \sim N(\mu, \sigma^2/n)$. ✓
- This result, which directly relates to Theorem 6.3, was derived on page 5 of Notes 8 via mgf's.

→ **EXAMPLE.** Wackerly 7, Exercise 7.11 (here, we assume $\sigma = 4$)

A forester discovered the average basal area follows normal distribution with $\sigma = 4$. If the forester samples $n = 9$ trees, find the probability that $\bar{Y}$ will be 2 square inches of population mean μ.

$Y_1 \cdots Y_n \sim N(\mu, \sigma^2)$ $\boxed{\sigma = 4}$ $n = 9$.

to compute values requires knowing σ ⚠

$P(|\bar{Y} - \mu| \le 2)$ $\quad \bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$.

$= P(\mu - 2 \le \bar{Y} \le \mu + 2)$ $\quad N(0,1)$.

$= P\left(\frac{-2}{\sqrt{\frac{\sigma^2}{n}}} \le \frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{n}}} \le \frac{2}{\sqrt{\frac{\sigma^2}{n}}}\right)$

recall $\Phi(x) + \Phi(-x) = 1$

$= \Phi\left(\frac{2}{\sqrt{\frac{\sigma^2}{n}}}\right) - \Phi\left(\frac{-2}{\sqrt{\frac{\sigma^2}{n}}}\right) = 2\Phi\left(\frac{2}{\sqrt{\frac{16}{9}}}\right) - 1$

$= \boxed{0.866}$

Fact: ① $Z \sim N(0,1)$ $\quad Z^2 \sim \chi^2(1)$

② $Z_i \sim N(0,1)$ independently. $\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$ ✓

$Z_i \sim N(0,1)$ $\quad N(\mu, \sigma^2)$

We have also seen Theorem 7.2 previously:

- If $Z_i = (Y_i - \mu)/\sigma$, then $\boxed{\sum_{i=1}^{n} Z_i^2}$ is distributed as a chi-square r.v. with n degrees of freedom (dof).
- This is a rephrasing of Theorem 6.4, except that here we assume $\mu_i = \mu\ \forall\, i$ and $\sigma_i = \sigma\ \forall\, i$.
- This result was also derived using mgf's, on page 5 of Notes 8.

$\boxed{2^\circ}$ Sampling distribution: $\sum_{i=1}^{n}\left(\frac{Y_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$

THE CHI-SQUARE DISTRIBUTION

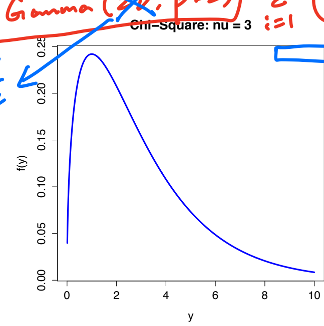
$\chi^2(\nu)$ same as Gamma $(\cancel{\alpha}, \beta = 2)$ $\quad \alpha = \frac{\nu}{2}$

NOTATION: $Y \sim \chi^2(\nu)$ $\quad \nu$: NUMBER OF DEGREES OF FREEDOM (DOF)

PDF: $f(y) = \dfrac{y^{\nu/2-1}e^{-y/2}}{2^{\nu/2}\Gamma(\nu/2)}$ $\quad y \in [0, \infty), \nu \in \mathbb{Z}^+$

EXPECTED VALUE: $E[Y] = \mu = \nu$ $\qquad$ VARIANCE: $V[Y] = \sigma^2 = 2\nu$

R FUNCTIONS: `dchisq(y,nu)` (PDF)
`pchisq(y,nu)` (CDF)
`qchisq(p,nu)` (INVERSE CDF)
`rchisq(k,nu)` (SIMULATION OF k χ^2 R.V.'S)



→ **EXAMPLE.** Wackerly 7, Exercise 7.23

7.23 Applet Exercise

a Use the applet *Chi-Square Probabilities and Quantiles* to find $P[Y > E(Y)]$ when Y has χ^2 distributions with 10, 40, and 80 df.

b What did you notice about $P[Y > E(Y)]$ as the number of degrees of freedom increases as in part (a)?

c How does what you observed in part (b) relate to the shapes of the χ^2 densities that you obtained in Exercise 7.22?

a) $P(Y > EY)$ $\qquad Y \sim \chi^2(\nu)$ $\quad E[Y] = \nu$

$= P(Y > \nu)$

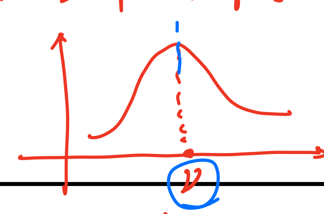
$= 1 - P(Y \le \nu)$

$= 1 - \text{pchisq}(\nu, \nu)$.

ν	$P(Y>\nu)$
10	0.440
40	0.470
80	0.479
⋮	⋮
∞	0.5

b) $\lim_{\nu \to \infty} P(Y > EY) = 0.5$

c). shape of pdf $\chi^2(\nu)$



$P(Y > \nu) \to \frac{1}{2}$

$\boxed{\nu \to \infty}$ shape becomes more and more like

why?

Why? Consequence of Central Limit theorem. (CLT). more ... a normal dist.

$\sigma^2 = \mathbb{E}[(Y_i-\mu)^2]$

good estimator of population variance.

Theorem 7.3 also references the chi-square distribution: if we are given n iid r.v.'s sampled from $N(\mu, \sigma^2)$, then

let $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ sample variance.

$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sim \chi^2(n-1)$$

3° Sampling distribution: $\frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 \sim \chi^2(n-1)$

is distributed as chi-square for $n-1$ dof. (Note that this uses the definition of sample variance given in Chapter 1.) The proof of this is long and involved; see page 358 of Wackerly 7 for the case $n = 2$.

Theorem 7.3 tells us how S^2 is distributed if we are dealing with iid data sampled from a normal distribution. It also allows us to specify the distribution from which $\sqrt{n}(\bar{Y} - \mu)/S$ is sampled, in situations where we wish to infer μ with σ being unknown:

need dist $\frac{\bar{Y}-\mu}{S/\sqrt{n}} = \sqrt{n}\cdot\frac{\bar{Y}-\mu}{\sigma} \sim N(0,1)$ $\quad \bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$ $\quad \frac{\bar{Y}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$\frac{S}{\sigma} = \sqrt{\frac{S^2(n-1)/\sigma^2}{n-1}} \sim \sqrt{\frac{\chi^2(n-1)}{n-1}}$ $\quad \frac{S}{\sigma} \sim \quad \frac{S^2\cdot(n-1)}{\sigma^2} \sim \chi^2(n-1).$ $\quad \frac{Z}{\sqrt{\frac{W}{\nu}}}$

$Z \sim N(0,1)$, $W \sim \chi^2(\nu)$, $W = \frac{(n-1)S^2}{\sigma^2}$. $\nu = n-1$

$t(\nu)$ is the ratio of $N(0,1)$ & independent square root $\sqrt{\frac{\chi^2(\nu)}{\nu}}$

Student's t Distribution

Notation: $Y \sim t(\nu)$ $\quad \nu$: number of degrees of freedom

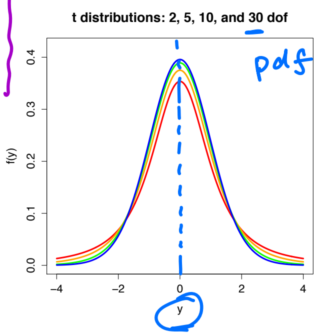
① PDF: $f(y) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1 + \frac{y^2}{\nu}\right)^{-\frac{\nu+1}{2}}$ $\quad y \in (-\infty, \infty), \nu \in \mathbb{Z}^+$

Expected Value: $E[Y] = \mu = 0$

Variance: $V[Y] = \sigma^2 = \nu/(\nu - 2)$ for $\nu > 2$

② $t(\nu) \sim \frac{N(0,1)}{\sqrt{\frac{\chi^2(\nu)}{\nu}}}$

R Functions:
```
dt(y,nu) (PDF)
pt(y,nu) (CDF)
qt(p,nu) (Inverse CDF)
rt(k,nu) (Simulation of k t r.v.'s)
```



To derive the pdf of the t distribution, one has to use methods outlined in Chapter 6, Section 6, which we will not cover. (Specifically, one has to determine a Jacobian to transform a bivariate distribution based on the combination of independent normal and χ^2 r.v.'s to a univariate t distribution.) The derivation is given in Casella & Berger, *Statistical Inference* (2nd ed.), on pages 223-224.

A sequence of t-distributed random variables T_1, T_2, ..., ordered by ascending sample size n, will, as $n \to \infty$, *converge in distribution* to the standard normal. (What this really means is that as $n \to \infty$, the difference between the true and estimated standard deviations goes to zero, i.e., $|S - \sigma| \to 0$, so one achieves nearly the same accuracy in probability calculations using either the t distribution or the standard normal distribution.) In practice, the rule of thumb for the amount of data you need to switch from using the t distribution to the standard normal distribution in $n \approx 30$, but this should always be confirmed via simulations.

$T_1 \sim t(n_1 - 1)$
$T_2 \sim t(n_2 - 1)$
$\vdots$
$n_1 < n_2 < n_3 \cdots$
$n \to \infty$

$T_1 \; T_2 \cdots T_k \cdots$
$\lim_{k\to\infty} T_k \to N(0,1).$

$t(\nu) = \frac{\sqrt{n}(\bar{Y}-\mu)}{S} \sim N(0,\sigma^2)$, $S \to \sigma$
$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

$t(\nu)$ when $\nu > 30$ then $t(\nu) \approx N(0,1)$.
(rule of Thumb).

4° sampling dist:

$\dfrac{\bar{Y}-\mu}{S/\sqrt{n}} \sim \underline{t(n-1)}$

$\sigma/\sqrt{n}$.

→ **EXAMPLE.** Wackerly 7, Exercise 7.11 (here, we assume $S = 4$)

A forester discovered the average basal area follows normal distribution ~~with $\sigma = 4$~~. If the forester samples $n=9$ trees, find the probability that $\bar{Y}$ will be 2 square inches of population mean μ.

$P(|\bar{Y}-\mu| \le 2)$.

$S=4$. $Y_1\ Y_2 \cdots Y_n \sim N(\mu, \sigma^2)$

$n=9$. $S=4$

$P(|\bar{Y}-\mu| \le 2) = P(\mu-2 \le \bar{Y} \le 2+\mu)$

$= P\left(\frac{-2}{S/\sqrt{n}} \le \boxed{\frac{\bar{Y}-\mu}{S/\sqrt{n}}} \le \frac{2}{S/\sqrt{n}}\right)$

$\sim t(n-1)$

$= P\left(-\frac{3}{2} \le \frac{\bar{Y}-\mu}{S/\sqrt{n}} \le \frac{3}{2}\right)$

$= pt\left(\frac{3}{2}, 8\right) - pt\left(-\frac{3}{2}, 8\right) = 0.828$

Another result that follows from Theorem 7.3 is the following:

new distribution.

If $W_1 \sim \chi^2(\nu_1)$ $W_2 \sim \chi^2(\nu_2)$ $W_1 \perp\!\!\!\perp W_2$ then: $\boxed{\dfrac{W_1/\nu_1}{W_2/\nu_2}} \sim F(\nu_1, \nu_2)$. definition

Suppose $Y_{11}\ Y_{12} \cdots Y_{1n_1} \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$. $(Y_{11} \cdots Y_{1n_1}) \perp\!\!\!\perp (Y_{21} \cdots Y_{2n_2})$.

$Y_{1i} \in \mathbb{R}$.

$Y_{21}\ Y_{22} \cdots Y_{2n_2} \overset{iid}{\sim} N(\mu_2, \sigma_2^2)$.

$\boxed{\dfrac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1-1)}$ where $S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_{1i}-\bar{Y}_1)^2$ sample variance.

$\boxed{\dfrac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_2-1)}$ where $S_2^2 =$ sample variance of Y_{2i}

$\sim \chi^2(n_1-1)$

5° sampling dist.

So: $\dfrac{\cancel{(n_1-1)}S_1^2/\sigma_1^2 \,/\, \cancel{(n_1-1)}}{\cancel{(n_2-1)}S_2^2/\sigma_2^2 \,/\, \cancel{(n_2-1)}} = \boxed{\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)}$

SNEDECOR'S F DISTRIBUTION

$F(4,6) \ne F(2,3)$

NOTATION: $Y \sim \mathrm{F}(\nu_1, \nu_2)$ ν_1, ν_2: NUMBER OF DOF

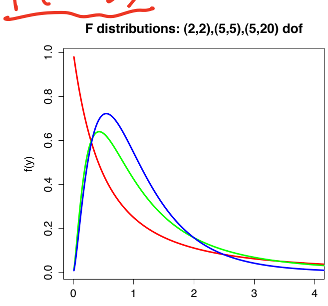
PDF: $f(y) = \dfrac{\Gamma[(\nu_1+\nu_2)/2]}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)}\left(\dfrac{\nu_1}{\nu_2}\right)^{\nu_1/2}\dfrac{y^{\nu_1/2-1}}{[1+(\nu_1/\nu_2)y]^{(\nu_1+\nu_2)/2}}$

$y \in [0, \infty), (\nu_1, \nu_2) \in \mathbb{Z}^+$

EXPECTED VALUE: $E[Y] = \mu = \nu_2/(\nu_2 - 2)$ FOR $\nu_2 > 2$

VARIANCE: $V[Y] = \sigma^2 = \dfrac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$ FOR $\nu_2 > 4$

R FUNCTIONS:
```
df(y,nu1,nu2) (PDF)
pf(y,nu1,nu2) (CDF)
qf(p,nu1,nu2) (INVERSE CDF)
rf(k,nu1,nu2) (SIMULATION OF k F R.V.'S)
```



Note that the derivation of the F distribution pdf is similar to that for the t distribution. More details are given on page 225 of Casella & Berger, where it is also called the *variance ratio* distribution.

→ **EXAMPLE.** Wackerly 7, Exercise 7.27

$\sigma_1^2 = \sigma_2^2$

$Y_{11}, Y_{12}, \dots, Y_{1n_1} \sim N(\mu_1, \sigma_1^2)$

7.27 **Applet Exercise** Refer to Example 7.7. If we take independent samples of sizes $n_1 = 6$ and $n_2 = 10$ from two normal populations with equal population variances, use the applet *F-Ratio Probabilities and Quantiles* to find

$Y_{21}, Y_{22}, \dots, Y_{2n_2} \sim N(\mu_2, \sigma_2^2)$

a $P(S_1^2/S_2^2 > 2)$.

b $P(S_1^2/S_2^2 < 0.5)$.

c the probability that one of the sample variances is at least twice as big as the other.

a) $P\left(\frac{S_1^2}{S_2^2} > 2\right) = P\left(\boxed{\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}} > 2\right) = 1 - pf(2, 5, 9) = 0.173$

$\sim F(\underbrace{n_1 - 1}_{5}, \underbrace{n_2 - 1}_{9})$

b) $P\left(\frac{S_1^2}{S_2^2} < \frac{1}{2}\right) = P\left(\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < \frac{1}{2}\right)$ $\quad F(5,9)$

$pf\left(\frac{1}{2}, 5, 9\right) = 0.23.$

c) P(one of sample variance is at least twice as big as the other)

$= P\left(\frac{S_1^2}{S_2^2} > 2 \text{ or } \frac{S_2^2}{S_1^2} > 2\right)$

$= P\left(\frac{S_1^2}{S_2^2} > 2\right) + P\left(\frac{S_2^2}{S_1^2} > 2\right)$

$= 0.173 + 0.23$

$= 0.403.$

→ **EXAMPLE.** Wackerly 7, Exercise 7.19

7.19 Ammeters produced by a manufacturer are marketed under the specification that the standard deviation of gauge readings is no larger than .2 amp. One of these ammeters was used to make ten independent readings on a test circuit with constant current. If the sample variance of these ten measurements is .065 and it is reasonable to assume that the readings are normally distributed, do the results suggest that the ammeter used does not meet the marketing specifications? [*Hint:* Find the approximate probability that the sample variance will exceed .065 if the true population variance is .04.]

$\sigma \le \sigma_{max} = 0.2 \text{(amp)}$

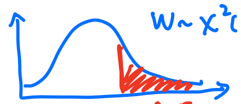
$Y_1 \dots Y_{10} \sim \boxed{N(\mu, \sigma^2)}.$

$S = 0.065.$

Whether this ammeter has std ≤ 0.2.
$\uparrow$
σ.

define $\boxed{W = \frac{(n-1)S^2}{\sigma^2}}$ $\quad S^2$: sample variance

$W \sim \chi^2(n-1) = \chi^2(9).$



observed.

if ammeter meets marketing specification, $\sigma \le 0.2$

observed $W = \frac{9 \times 0.065}{\sigma^2} \ge \frac{9 \times 0.065}{(0.2)^2} = 14.625$

$\boxed{P(W \ge 14.625)}.$
$\chi^2(9)$ $= 1 - pchisq(14.625, 9)$

$= 0.102$

prob of ammeter meets MS is 0.102

Standard judgement

prob $\le 0.05 \Rightarrow$ reject

(more of this type of reasoning in Stat 431).

→ **EXAMPLE.** Wackerly 7, Exercise 7.33

7.33 Use the structures of T and F given in Definitions 7.2 and 7.3, respectively, to argue that if T has a t distribution with ν df, then $U = T^2$ has an F distribution with 1 numerator degree of freedom and ν denominator degrees of freedom.

$t(\nu) \sim T = \frac{Z}{\sqrt{\frac{W}{\nu}}}$ $\quad Z \sim N(0,1)$ $\quad W \sim \chi^2(\nu)$

$T^2 = \frac{Z^2}{\frac{W}{\nu}} = \frac{Z^2/1}{W/\nu} \sim F(1, \nu).$ $\quad Z^2 \sim \chi^2(1)$, $W \sim \chi^2(\nu)$

check defn for F-dist.

→ **EXAMPLE.** Wackerly 7, Exercise 7.37

7.37 Let $Y_1, Y_2, \dots, Y_5$ be a random sample of size 5 from a normal population with mean 0 and variance 1 and let $\bar{Y} = (1/5)\sum_{i=1}^5 Y_i$. Let Y_6 be another independent observation from the same population. What is the distribution of

$N(0,1)$ $\quad Y_6 \sim N(0,1).$

a $W = \sum_{i=1}^5 Y_i^2$? Why?

b $U = \sum_{i=1}^5 (Y_i - \bar{Y})^2$? Why?

c $\sum_{i=1}^5 (Y_i - \bar{Y})^2 + Y_6^2$? Why?

a) $W = \sum_{i=1}^5 Y_i^2 \sim \chi^2(5).$ $\quad Y_i \sim N(0,1)$ independent. $Y_i^2 \sim \chi^2(1)$.

b) $U = \sum_{i=1}^5 (Y_i - \bar{Y})^2 \sim \chi^2(n-1) = \chi^2(4).$

$= 4S^2 = \frac{(n-1) \cdot S^2}{\sigma^2}$ $\quad \sigma^2 = 1$

$S^2 = \frac{1}{5-1} \sum_{i=1}^5 (Y_i - \bar{Y})^2.$

c) $\sum_{i=1}^5 (Y_i - \bar{Y})^2 + Y_6^2 = U + Y_6^2 \sim \chi^2(5)$

arbitrary.

And now we move into Section 7.4 and consider Theorem 7.4, the *central limit theorem*:

$Y_1 \cdots Y_n \overset{iid}{\sim} P$ some dist P). mean μ Variance σ^2.

"loosely speaking" for large n, $\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$.

RoT $n \geq 30$.

CLT: $\lim_{n\to\infty} P\left(\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x)$ ← cdf for $N(0,1)$.

Proving Theorem 7.4 involves the use of mgf's:

Pf: let $V_i = \frac{Y_i - \mu}{\sigma}$ $\quad E[V_i] = 0 \quad V[V_i] = 1.$

$$X = \frac{\bar{Y}-\mu}{\sigma/\sqrt{n}} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_i = \sum_{i=1}^{n} a_i V_i \qquad a_i = \frac{1}{\sqrt{n}}.$$

to find dist of X. V_i has the same dist.

μ_k' k-moment of V_i

$$m_X(t) = \prod_{i=1}^{n} m_{V_i}\left(\frac{1}{\sqrt{n}}t\right) = \left(m_{V_1}\left(\frac{1}{\sqrt{n}}t\right)\right)^n$$

$$= \left(1 + \frac{t}{\sqrt{n}}\cdot \underbrace{\mu_1'}_{0} + \left(\frac{t}{\sqrt{n}}\right)^2 \cdot \frac{\mu_2'}{2!} + \cdots\right)^n$$

$$= \left(1 + \frac{t^2}{2n} + [\cdots]\right)^n \to \lim_{n\to\infty}\left(1+\frac{t^2}{2n}\right)^n = e^{\frac{t^2}{2}}$$

Contains $\frac{1}{n^{1.5}}$ $\frac{1}{n^2}$ ...

when n is large, they are small compared to $\frac{1}{n}$.

$m_X(t) = e^{\frac{t^2}{2}} \Rightarrow X \sim N(0,1)$

$\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a$

→ **EXAMPLE.** Wackerly 7, Exercise 7.45

7.45 Workers employed in a large service industry have an average wage of \$7.00 per hour with a standard deviation of \$.50. The industry has 64 workers of a certain ethnic group. These workers have an average wage of \$6.90 per hour. Is it reasonable to assume that the wage rate of the ethnic group is equivalent to that of a random sample of workers from those employed in the service industry? [*Hint:* Calculate the probability of obtaining a sample mean less than or equal to \$6.90 per hour.]

$Y_i \sim P$ $\quad \mu = 7 \quad \sigma = 0.5$ $\quad \bar{Y} = 6.90$

$P(\bar{Y} \leq 6.9) = P\left(\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}} \leq \frac{6.9-\mu}{\sigma/\sqrt{n}}\right)$ r.v

by CLT. $n = 64$ $\sim N(0,1)$

$\approx \Phi\left(\frac{6.9-7}{0.5/\sqrt{64}}\right) = \Phi(-1.6) = 0.054.$

compare with 0.05.

→ **EXAMPLE.** Wackerly 7, Exercise 7.49

7.49 The length of time required for the periodic maintenance of an automobile or another machine usually has a mound-shaped probability distribution. Because some occasional long service times will occur, the distribution tends to be skewed to the right. Suppose that the length of time required to run a 5000-mile check and to service an automobile has mean 1.4 hours and standard deviation .7 hour. Suppose also that the service department plans to service 50 automobiles per 8-hour day and that, in order to do so, it can spend a maximum average service time of only 1.6 hours per automobile. On what proportion of all workdays will the service department have to work overtime?

$\mu = 1.4 \quad \sigma = 0.7. \quad n = 50$

cdf for $N(0,1)$

$$P(\bar{Y} > 1.6) = P\left(\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}} > \frac{1.6-\mu}{\sigma/\sqrt{n}}\right) \approx 1 - \Phi\left(\frac{1.6-1.4}{0.7/\sqrt{50}}\right) = 0.022.$$

$n = 50 > 30$

use CLT. $\frac{\bar{Y}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

We conclude this set of notes by mentioning two concepts that are associated with the central limit theorem.

Law of Large numbers (LLN)

Let $X_1, X_2, \ldots,$ be iid random variables, with mean μ and variance $\sigma^2 < \infty$ (i.e., finite variance[a]). Let $\bar{X} = (1/n)\sum_{i=1}^{n} X_i$. Then, for every $\epsilon > 0$, we have that

$$\boxed{\lim_{n\to\infty} P(|\bar{X}_n - \mu| \geq \varepsilon) = 0} \quad \text{or} \quad \bar{X}_n \xrightarrow{P} \mu$$

This is the $\boxed{\textit{weak law of large numbers.}}$ This differs from the CLT in that here, the sample mean $\bar{X}$ "converges in probability" to μ. It says nothing about the *distribution* of $\bar{X}$. (In 226-speak, the weak law says that $\bar{X}$ is a *consistent* estimator of μ.)

The $\boxed{\textit{strong law of large numbers}}$ tweaks the weak law:

$$P(\lim_{n\to\infty} \bar{X}_n = \mu) = 1 \quad \text{or} \quad \bar{X}_n \xrightarrow{a.s.} \mu$$

$$\text{or } P(\lim_{n\to\infty} |\bar{X}_n - \mu| \geq \varepsilon) = 0$$

Instead of saying that $\bar{X}$ "converges in probability" to μ (weak law), it says that $\bar{X}$ "converges almost surely" to μ (which is a stronger statement).

What, effectively, is the difference between these two laws?

intepretation of weak LLN → go back to sample
strong LLN space!

⇓

advanced prob class.

---

[a]counterexample: the Cauchy distribution...