United States Census Bureau

South Dakota State University

SDSU

South Dakota Census Data Center

Challenges in the data

Margins of error, systematic errors, and

The American Community Survey

- The ACS is much more timely than the long form was
- The Margin of Error in the ACS is OK for Sioux Falls and Rapid City – in smaller communities, we really need to look at, and understand, the potential for error.

Subject	Murdo city, South Dakota			
	Estimate	Margin of Error	Percent	Percent Margin of Error
HOUSEHOLDS BY TYPE				
Total households	223	+/-33	223	(X)
Family households (families)	124	+/-35	55.6%	+/-11.7
With own children under 18 years	65	+/-25	29.1%	+/-9.6
Married-couple family	88	+/-29	39.5%	+/-11.1
With own children under 18 years	33	+/-18	14.8%	+/-7.6
Male householder, no wife present, family	9	+/-12	4.0%	+/-5.4
With own children under 18 years	9	+/-12	4.0%	+/-5.4
Female householder, no husband present, family	27	+/-15	12.1%	+/-6.7
With own children under 18 years	23	+/-13	10.3%	+/-5.8
Nonfamily households	99	+/-26	44.4%	+/-11.7
Householder living alone	93	+/-26	41.7%	+/-12.0
65 years and over	56	+/-23	25.1%	+/-10.6
Households with one or more people under 18 years	65	+/-25	29.1%	+/-9.6
Households with one or more people 65 years and over	81	+/-20	36.3%	+/-9.1
Average household size	2.43	+/-0.40	(X)	(X)
Average family size	3.40	+/-0.76	(X)	(X)
RELATIONSHIP				
Population in households	542	+/-103	542	(X)
Householder	223	+/-33	41.1%	+/-6.8
Spouse	87	+/-27	16.1%	+/-4.3
Child	205	+/-77	37.8%	+/-9.1
Other relatives	6	+/-9	1.1%	+/-1.7
Nonrelatives	21	+/-19	3.9%	+/-3.4
Unmarried partner	7	+/-8	1.3%	+/-1.5

Subject	Yankton city, South Dakota			
	Estimate	Margin of Error	Percent	Percent Margin of Error
HOUSEHOLDS BY TYPE				
Total households	5,780	+/-254	5,780	(X)
Family households (families)	3,403	+/-225	58.9%	+/-3.2
With own children under 18 years	1,482	+/-135	25.6%	+/-2.4
Married-couple family	2,678	+/-199	46.3%	+/-3.1
With own children under 18 years	1,027	+/-115	17.8%	+/-2.2
Male householder, no wife present, family	143	+/-66	2.5%	+/-1.1
With own children under 18 years	21	+/-25	0.4%	+/-0.4
Female householder, no husband present, family	582	+/-145	10.1%	+/-2.4
With own children under 18 years	434	+/-134	7.5%	+/-2.3
Nonfamily households	2,377	+/-224	41.1%	+/-3.2
Householder living alone	1,987	+/-224	34.4%	+/-3.5
65 years and over	871	+/-156	15.1%	+/-2.6
Households with one or more people under 18 years	1,589	+/-134	27.5%	+/-2.4
Households with one or more people 65 years and over	1,700	+/-161	29.4%	+/-2.5
Average household size	2.18	+/-0.07	(X)	(X)
Average family size	2.83	+/-0.14	(X)	(X)

Subject	Sioux Falls, SD Metro Area			
	Estimate	Margin of Error	Percent	Percent Margin of Error
HOUSEHOLDS BY TYPE				
Total households	88,307	+/-733	88,307	(X)
Family households (families)	59,182	+/-831	67.0%	+/-0.9
With own children under 18 years	29,783	+/-694	33.7%	+/-0.7
Married-couple family	46,497	+/-909	52.7%	+/-1.0
With own children under 18 years	21,092	+/-671	23.9%	+/-0.7
Male householder, no wife present, family	3,280	+/-336	3.7%	+/-0.4
With own children under 18 years	1,840	+/-310	2.1%	+/-0.3
Female householder, no husband present, family	9,405	+/-522	10.7%	+/-0.6
With own children under 18 years	6,851	+/-467	7.8%	+/-0.5
Nonfamily households	29,125	+/-941	33.0%	+/-0.9
Householder living alone	23,576	+/-856	26.7%	+/-0.9
65 years and over	7,404	+/-385	8.4%	+/-0.4
Households with one or more people under 18 years	31,278	+/-664	35.4%	+/-0.7
Households with one or more people 65 years and over	17,042	+/-339	19.3%	+/-0.4
Average household size	2.48	+/-0.02	(X)	(X)
Average family size	3.01	+/-0.03	(X)	(X)

Systematic Errors

Random errors can be handled statistically

- **Systematic errors**
  - hard to detect
  - defy statistical analysis because all of the data is off in the same direction (either to high or too low).
  - finding and correcting systematic errors requires care and time.

Examples of Systematic Errors

A broken chain always measures long

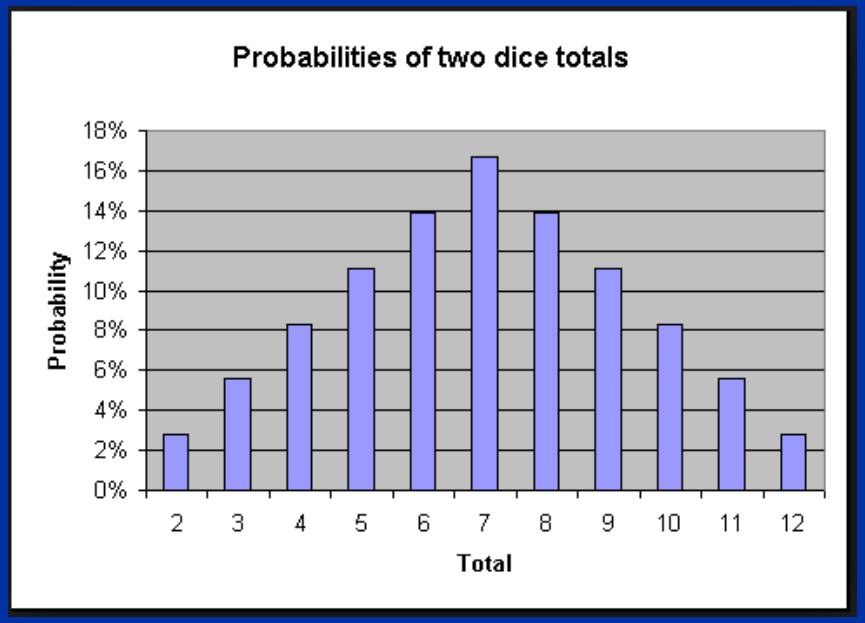
- Reservation/Indian Majority County Sampling & surveys
- Hutterite Populations

A stretched chain always measures short

- Larger families, multi-family residences and addresses than their white neighbors
- Faulk County's population is 20% Hutterites, living on Colony – this affects things like median household income, etc.

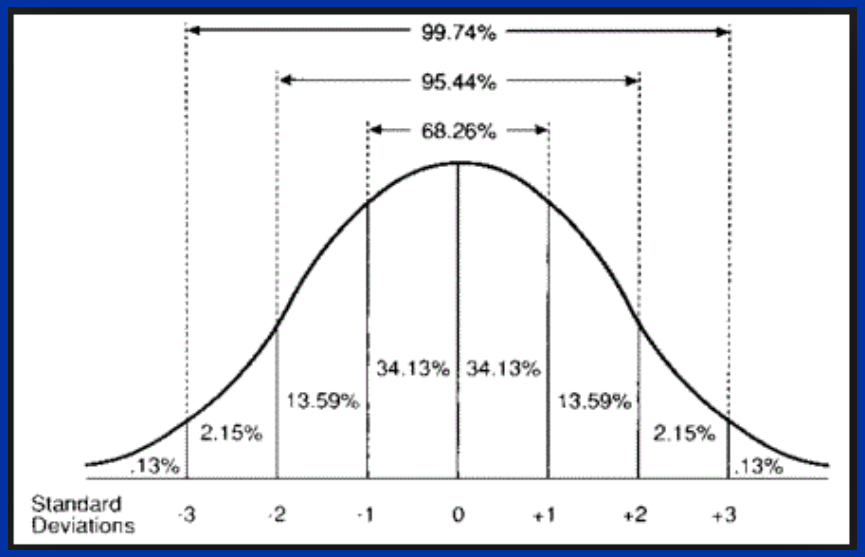
Statistics started with gamblers

- Note that the probabilities of two dice totals – as occur when shooting craps – strongly resembles a normal or "bell" curve



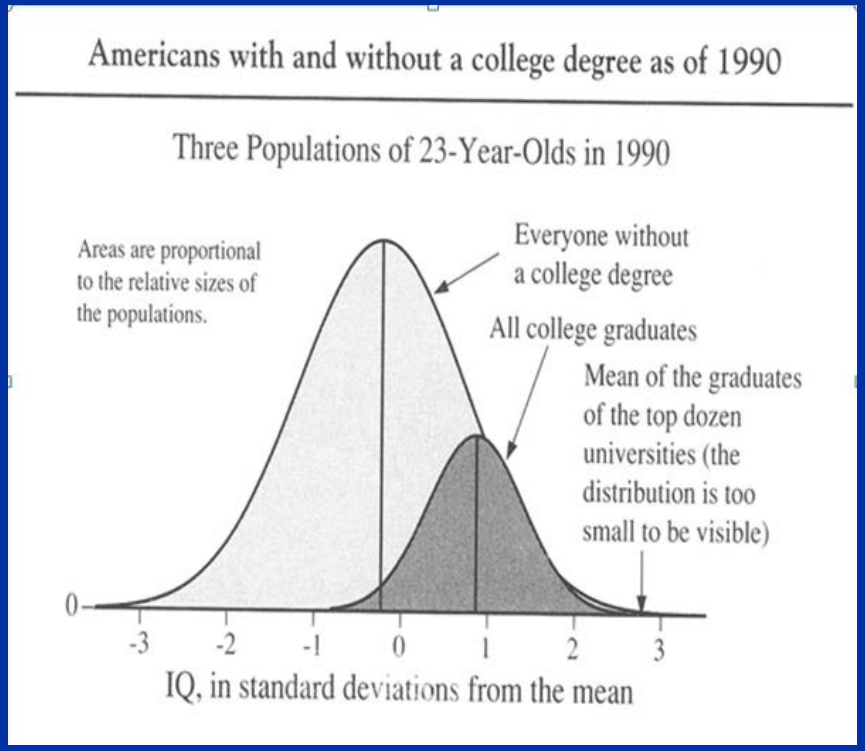
Then mathematicians got involved

- The Bell curve is symmetric, smooth, precise – all in all very appealing.



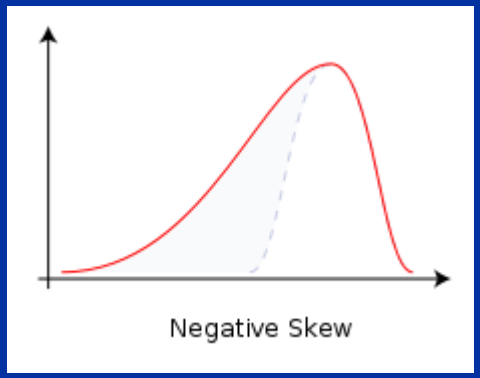
As the Curve Skews . . .

- Most of our statistical analysis & corrections are based on the Bell curve . . . The "normal" curve.
- In demography, the curve is often skewed to one side or the other, requiring different calculations.



The benefits of skewed curves

More Low Values



More High Values

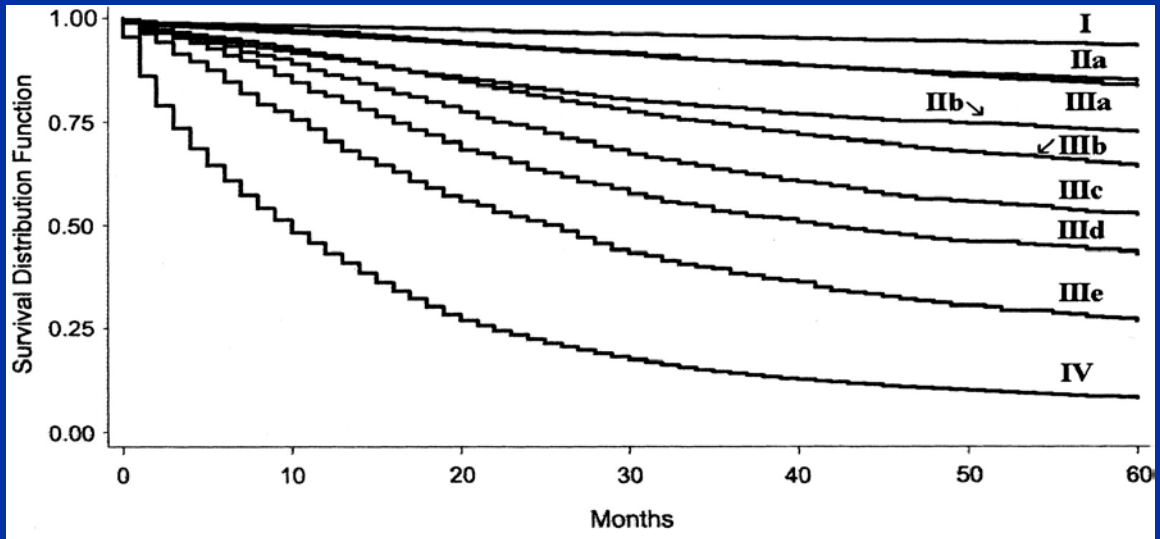
Skewed isn't bad

- Positively skewed – the shortest possible lifespan is 0.
- Data with a lower limit will be positively skewed, Data with a high limit will be negatively skewed
- Hutterite colonies skew median household income negatively
- Life expectancy
- There is a pattern
- Outliers can cause skewed data

Skewed data can show systematic errors

- Negative Skew on Reservation or Indian Majority Counties
- Negative skew with larger household sizes
- Positive skew with increasing age of householder
- Median Per Capita or Household Income
- Home Ownership

The medical folks make curves, too
Survival by American Joint Committee on Cancer sixth edition staging with proposed lymph node (N) stages. (colon cancer)



Stage	0 mo		30 mo			60 mo		
	Survival, %	No.	Survival, %	No.	*P**	Survival, %	No.	*P**
I	100	14500	96.1	8581	—	93.2	4514	—
IIa	100	28535	91.0	2105	<.001	84.7	8494	<.001
IIb	100	5826	80.2	3060	<.001†	72.2	1611	<.001†
IIIa	100	1989	91.4	1120	NS‡	83.4	551	NS‡
IIIb	100	15946	77.3	7786	<.001§	64.1	3579	<.001§
IIIc	100	4092	67.1	1697	<.001	52.3	725	<.001
IIId	100	2655	57.3	908	<.001	43.0	384	<.001
IIIe	100	1853	43.1	434	<.001	26.8	141	<.001
IV	100	20802	17.3	1832	<.001	8.1	432	<.001

JNCI