Markov's Inequality

Jonathan Auerbach

9/28/2018

Percentiles are often used to summarize distributions such as income.

- The 90-percentile, p_{90}, is the income such that 90% of households make less than p_{90}.
- i.e. randomly select a houshold in the United States and record the income as X. If done many times: $\mathbb{P}(X < p_{90}) \approx .9$.
- $p_{.90}$ is also called the upper decile or the .9-quantile.

What is 90-percentile household income in the United States?

- Household income includes salaries, unemployment insurance, disability or child support payments, and personal, business, or investment income.
- It includes all residents in a house over the age of 15.

What is 90-percentile household income in the United States?

- Household income includes salaries, unemployment insurance, disability or child support payments, and personal, business, or investment income.
- It includes all residents in a house over the age of 15.
- For reference, the average household income, μ is 70,000. Does this matter?

Markov's Inequality connects the mean and percentile of any population.

- $\mu \geq p \cdot \mathbb{P}(X > p)$ provided both X and $p \geq 0$

$$
\begin{aligned}
\mu &= \sum_{x=0}^{\infty} x \cdot \mathbb{P}(X = x) \\
&= \sum_{x=0}^{p} x \cdot \mathbb{P}(X = x) + \sum_{x=p+1}^{\infty} x \cdot \mathbb{P}(X = x) \\
&\geq \sum_{x=p+1}^{\infty} p \cdot \mathbb{P}(X = x) \\
&= p \cdot \sum_{x=p+1}^{\infty} \mathbb{P}(X = x) \\
&= p \cdot \mathbb{P}(X > p)
\end{aligned}
$$

Markov's Inequality $\mu \geq p_{90} \cdot \mathbb{P}(X > p_{90})$

- Rearranging terms:
  $.1 = \mathbb{P}(X > p_{90}) \leq \frac{\mu}{p_{90}} \Rightarrow p_{90} \leq \frac{\mu}{.1} = 10 \cdot \mu$
- So an average income of 70,000 implies a 90-percentile income of no more than $10 \cdot 70,000 = 700,000$
- We can bound the percentile just by knowing the mean!
- This is useful for many problems.

What is the probability the average of n household incomes, randomly selected from the population, is different than the population average?

- Let $Y = |\frac{1}{n}\sum_{i=1}^{n} X_i - \mu|$
- It turns out, $\mu_y \underset{n\to\infty}{\to} 0$, where μ_y is the average of Y.
- Therefore, $\mathbb{P}(Y > p) \leq \frac{\mu_y}{p} \underset{n\to\infty}{\to} 0$
- This is known as the Weak Law of Large Numbers.

References

1.