

Using Statistics to Evaluate Healthy Lifestyles and Lottery Fairness

Posted on October 16, 2018

Introduction

Statistical studies often begin with questions that appear simple but require careful analysis. Two useful examples are “Is your life healthy?” and “Are lotteries fair?” The first question requires researchers to collect information about health behaviors, choose an appropriate sample, measure relevant variables, and determine whether the results represent a wider population. The second requires an understanding of probability, expected value, public revenue, consumer behavior, and the different meanings of fairness.

Lottery advertisements often make winning appear to be a realistic way to become rich quickly. However, the probability of winning a major jackpot is extremely small. A lottery may use a fair random drawing while still offering participants an unfavorable financial return. Similarly, a health survey may produce percentages that appear authoritative, but the findings are useful only when the sample, questions, measurements, and limitations are properly understood.

These two topics demonstrate why statistical literacy matters in everyday life. Numbers do not interpret themselves. People must ask how data were obtained, which population they represent, what probabilities mean, whether important information has been omitted, and whether a conclusion is supported by the available evidence.

Why Statistical Reasoning Matters

Statistics involves more than collecting numbers or calculating percentages. It provides methods for gathering evidence, describing patterns, estimating unknown characteristics, and evaluating uncertainty. Bennett et al. (2018) explain that statistical reasoning allows people to assess claims encountered in areas such as health, politics, business, education, and personal finance.

Without statistical reasoning, a person may be influenced by a dramatic testimonial, a large advertised jackpot, or an isolated health claim. A lottery winner may appear in an advertisement, but the millions of people who bought losing tickets remain invisible. A health report may state that one group experienced fewer illnesses, but the difference may be associated with age, income, education, smoking, healthcare access, or another factor.

Good statistical reasoning requires attention to both the evidence presented and the information left

out. A convincing analysis should explain the population, sample, variables, methods, uncertainty, and possible alternative explanations.

Are Lotteries Fair?

The fairness of a lottery cannot be answered with a simple yes or no. It depends on what is meant by fairness.

A lottery may be procedurally fair when every valid ticket has the same chance of winning, the drawing is random, the rules are applied consistently, and the results are independently verified. However, equal chances do not mean that the game offers a favorable financial investment. A lottery can be fair in the way it selects a winner while being financially disadvantageous to almost every participant.

This distinction is essential. The question is not only whether the drawing process is honest. It is also whether players understand the probability of winning, the likely financial return, the use of lottery revenue, and the risks associated with repeated play.

Lottery Probability and the Chance of Winning

Probability measures how likely an event is to occur. It can be expressed as a fraction, decimal, percentage, or statement of odds. A probability close to zero describes an event that is possible but highly unlikely.

For example, the official odds of winning the Powerball jackpot are approximately 1 in 292.2 million. The overall odds of winning any prize are much better, approximately 1 in 24.87, but most non-jackpot prizes are relatively small (Multi-State Lottery Association, n.d.). The difference between winning “a prize” and winning the advertised jackpot must therefore be made clear.

Lottery advertisements usually emphasize the possible reward rather than the overwhelming probability of losing. They may show a large jackpot, an excited winner, or an image of financial freedom. These presentations are emotionally powerful because people tend to focus on vivid possibilities rather than base rates.

The existence of a winner does not demonstrate that purchasing a ticket was statistically advantageous. Someone must eventually win when enough tickets are sold and drawings continue, but the probability for each individual ticket remains extremely small. Buying more tickets increases the chance of winning, but it also increases the amount of money at risk. It does not transform the lottery into a reliable investment strategy.

The Difference Between Odds and Expected Value

Probability tells a player how likely a particular outcome is. Expected value estimates the average financial result that would occur over a very large number of similar plays.

The expected value of a lottery ticket is calculated by multiplying each possible prize by the probability of winning it, adding those products together, and subtracting the ticket price. Taxes, shared jackpots, payment options, and the possibility that no jackpot is won may also affect a more detailed calculation.

State lotteries are designed so that ticket sales exceed the total amount returned to players over time. Some money is used for prizes, some supports operating expenses and retailers, and the remaining net revenue is transferred to government programs. Therefore, the average player must lose money over repeated participation.

This does not mean that no person can receive more than the cost of a ticket. Individual winners clearly do. Expected value describes the long-run average across all tickets rather than the guaranteed outcome for one person.

A participant may still purchase a ticket for entertainment. Spending a small amount for the excitement of imagining a possible win is different from treating lottery tickets as a financial plan. The statistical problem arises when people misunderstand low probabilities, underestimate repeated losses, or believe that a jackpot is likely to solve their financial difficulties.

Does Voluntary Participation Make a Lottery Fair?

The original essay correctly observes that lottery participation is generally voluntary. Unlike an ordinary tax, no person is legally required to purchase a ticket. Supporters therefore argue that lottery revenue is preferable to compulsory taxation because people choose whether to contribute.

Voluntary participation is relevant, but it does not settle the fairness question. A choice is fully informed only when consumers understand the rules, odds, likely returns, and risks. A technically voluntary decision may still raise ethical concerns when advertising encourages unrealistic expectations or when participants have difficulty interpreting very small probabilities.

Behavioral research suggests that consumers do not always evaluate lottery risks according to conventional models of rational financial decision-making. Kearney (2005) found that lottery demand raised questions about how accurately consumers perceived risks and returns. More recent research has also associated lottery spending with weak statistical reasoning and the tendency to give excessive decision weight to very small probabilities (Lockwood et al., 2021).

Therefore, voluntariness should be considered alongside transparency and informed decision-making.

A fair system should not conceal the odds, misrepresent probable returns, or deliberately encourage vulnerable players to believe that regular participation is a realistic path out of poverty.

Do Lotteries Lower Taxes?

The original essay states that lotteries help lower tax rates because they produce billions of dollars in revenue that governments can use for education and public services. This claim requires qualification.

State lotteries do generate substantial government revenue. According to the U.S. Census Bureau (2026), state lottery ticket sales reached approximately \$104.7 billion in fiscal year 2024. About \$70.2 billion was paid in prizes, leaving \$34.5 billion in net lottery revenue before considering how individual states allocated the funds. Net revenue represented roughly 33% of ticket sales.

However, lottery revenue does not automatically reduce tax rates. Whether taxes are reduced depends on legislative decisions, budget needs, economic conditions, and the way lottery proceeds are incorporated into state finances. A government may use lottery revenue to supplement existing spending, replace part of another funding source, or support a designated program without changing taxes at all.

Many states direct at least some lottery revenue toward education, parks, veterans' programs, or other public services. These expenditures may benefit the public, but the existence of a socially useful destination does not prove that lottery funding is the fairest or most reliable way to finance it.

Lottery revenue can also fluctuate according to ticket sales and jackpot sizes. Essential public services require predictable funding, whereas lottery receipts may vary from year to year. Governments should therefore avoid presenting lotteries as a complete substitute for stable and transparent public finance.

The Distributional Fairness of Lotteries

Another important question is who bears the financial burden of lottery participation. A lottery may give every ticket an equal chance, but its economic effects may not be distributed equally across society.

Spending ten dollars on lottery tickets has a very different effect on a household with a low income than on a wealthy household. Even when people in different income groups spend similar dollar amounts, the purchase consumes a larger proportion of the lower-income household's available resources.

Research has frequently characterized state lotteries as regressive because spending does not rise in proportion to income and may impose a larger relative burden on people with fewer financial

resources (Carruthers & Wanamaker, 2020). The degree of regressivity can vary by game, location, jackpot size, and player behavior, so it should not be assumed that every lottery affects every group identically.

This issue complicates the claim that lotteries provide painless public revenue. The government receives money voluntarily, but some of that revenue may come from people who can least afford repeated losses. The benefits of lottery-funded programs may also be distributed differently from the costs.

A complete evaluation of fairness must therefore examine more than random selection. It must consider advertising, consumer understanding, income differences, problem gambling, the use of proceeds, and whether the system places disproportionate costs on vulnerable groups.

Honest Advertising and Responsible Participation

Lottery advertising should present prizes without creating misleading impressions about the chance of winning. The jackpot amount may be accurate, yet an advertisement can still be misleading when it gives extraordinary attention to winners while minimizing the probability of loss.

Responsible communication should clearly disclose the odds and avoid language suggesting that winning is a realistic financial strategy. Consumers should also recognize that previous results do not change the probability of a future independent drawing. A number that has not appeared recently is not automatically “due,” and a ticket does not become more likely to win because the purchaser has lost repeatedly.

Lottery tickets should be purchased, when legally permitted, only with money a person can afford to lose. They should not replace savings, debt repayment, insurance, retirement planning, or other financial priorities.

Is Your Life Healthy?

The question “Is your life healthy?” also requires statistical reasoning. Health is multidimensional and cannot be measured accurately through one behavior or one survey response.

A healthy lifestyle may involve nutrition, physical activity, tobacco avoidance, sleep, stress management, alcohol consumption, preventive healthcare, social relationships, and mental well-being. A person may perform well in some areas and poorly in others. Therefore, researchers must define what they mean by “healthy” before collecting and interpreting data.

The original essay refers to a study in Colorado. A state health study can be used to illustrate the distinction among a population, sample, variable, and statistic. It also demonstrates why percentages

should not be accepted without investigating how the respondents were selected and how the questions were asked.

Population and Sample in a Health Study

The population is the complete group about which researchers want to draw a conclusion. If a Colorado study aims to describe the health behaviors of all adults living in the state, then all Colorado adults form the target population.

Researchers usually cannot question every member of such a large population. Instead, they collect information from a sample. A sample is the smaller group actually observed or surveyed.

A numerical result calculated from the sample is called a statistic. For example, the percentage of surveyed adults who exercise regularly is a sample statistic. The corresponding but usually unknown percentage among all Colorado adults is a population parameter.

The purpose of statistical inference is to use the sample statistic to estimate the population parameter. This process is valid only when the sample and method provide adequate evidence.

A large sample is useful, but size alone does not guarantee accuracy. Thousands of respondents selected through a biased process may provide a less reliable estimate than a smaller, well-designed probability sample.

Representativeness of the Health Sample

A representative sample resembles the target population in the characteristics relevant to the research question. Adults of different ages, geographic locations, income levels, educational backgrounds, racial and ethnic groups, and health conditions should have an appropriate opportunity to be included.

Selection bias occurs when the method of recruitment systematically excludes or overrepresents certain groups. For example, an online health survey may underrepresent people with limited internet access. A telephone survey may miss people who do not answer unknown numbers. A survey conducted only in fitness centers would likely overrepresent adults who are already physically active.

Nonresponse can create another problem. People who choose to answer health questions may differ from those who refuse. Health-conscious individuals might be more willing to participate, while people with serious health problems or limited time might be less likely to respond.

Researchers may use weighting to adjust the sample so that it better reflects known population characteristics. Weighting can improve estimates, but it cannot completely correct every form of selection or measurement error.

Measuring a Healthy Lifestyle

Health variables must be defined in measurable terms. A survey cannot rely only on the vague question, “Do you live a healthy life?” Respondents may interpret “healthy” differently, and some may provide socially desirable rather than accurate answers.

Researchers can ask more specific questions about the frequency, duration, or quantity of behaviors. These might include weekly physical activity, daily fruit and vegetable consumption, smoking status, average sleep duration, preventive medical visits, or alcohol intake.

Even specific questions may contain measurement error. People may forget what they ate, overestimate their exercise, underestimate alcohol use, or report their current weight inaccurately. Self-reported information is valuable, particularly in large studies, but researchers must recognize its limitations.

Objective measurements such as accelerometer data, laboratory tests, medical records, blood pressure readings, or directly measured height and weight may improve accuracy. However, they can also be more expensive, intrusive, and difficult to collect from large populations.

A strong study selects measures that are reliable, valid, appropriate to the population, and closely connected to the health outcome being investigated.

Lessons From the Nurses’ Health Study

Long-term cohort studies demonstrate how researchers can examine associations between lifestyle and health. The original Nurses’ Health Study began in 1976 and enrolled 121,700 registered nurses. Participants have been followed through repeated questionnaires, allowing researchers to study changes in behavior and health over time (Nurses’ Health Study, n.d.).

Research using this cohort found that combinations of behaviors involving diet, physical activity, smoking avoidance, moderate alcohol consumption, and healthy body weight were associated with substantially lower coronary heart disease risk among women (Stampfer et al., 2000).

The study provides stronger evidence than a one-time opinion poll because researchers can measure lifestyle factors before many health outcomes develop and continue following participants over time. Nevertheless, it remains an observational study. Researchers observe what participants do rather than randomly assigning them to smoke, exercise, gain weight, or consume particular diets for decades.

The findings can therefore identify important associations, but causal interpretation still requires consideration of confounding, measurement error, biological plausibility, consistency with other studies, and alternative explanations.

Correlation Does Not Automatically Establish Causation

Suppose a health survey finds that physically active people report better health. This is a correlation: two variables are associated.

The association may partly reflect a causal effect of exercise on health. However, other explanations are possible. Healthier people may be more able to exercise, producing reverse causation. Active participants may also differ in income, diet, age, education, healthcare access, or smoking behavior.

A confounding variable is a third factor associated with both the explanatory variable and the outcome. Researchers use study design and statistical adjustment to reduce confounding, but adjustment is limited to factors that were measured accurately and included in the analysis.

Randomized controlled trials provide stronger protection against confounding because random assignment tends to distribute known and unknown characteristics across groups. However, many long-term lifestyle questions cannot be studied through random assignment for practical or ethical reasons. Observational evidence therefore remains essential in public health.

Readers should avoid interpreting every association as proof of cause, but they should not conclude that observational research has no value. Strong conclusions often emerge from multiple studies using different designs and producing consistent findings.

Statistical Significance and Practical Importance

Health studies frequently report that a result is statistically significant. Statistical significance generally indicates that the observed pattern would be relatively unlikely under a specified null hypothesis if the study assumptions were correct.

It does not reveal whether an effect is large, important, clinically meaningful, or free from bias. A very large sample may identify a tiny difference as statistically significant, while a small study may fail to detect an important effect because it lacks statistical power.

Readers should examine effect sizes and confidence intervals rather than relying only on a p-value. An effect size describes the magnitude of a difference or association. A confidence interval communicates the range of estimates reasonably compatible with the data under the method's assumptions.

For an individual deciding whether a lifestyle is healthy, practical importance matters greatly. A statistically detectable difference that has little effect on illness, functioning, or quality of life may not justify a major behavioral change. Conversely, a well-established behavior such as avoiding tobacco

can have substantial health importance even though an individual's exact future risk cannot be predicted with certainty.

Generalizing Health Findings

A study's results should be generalized only to populations reasonably represented by its participants.

A Colorado survey may provide valuable evidence about adults in that state, but its findings may not apply unchanged to every country or population. Climate, culture, income, healthcare systems, food availability, age distribution, and other conditions can influence health behaviors.

Similarly, the Nurses' Health Study has produced highly valuable evidence, but the original participants were female registered nurses within specified age ranges. Their education, occupation, and access to healthcare may differ from those of the general population. Researchers must consider whether a particular association is likely to apply to men, younger people, different occupational groups, or populations outside the United States.

Generalizability is not an all-or-nothing property. A study may provide strong biological or behavioral insight while still requiring replication in more diverse groups.

Shared Statistical Lessons From Both Questions

The lottery and health examples appear unrelated, but they require several of the same reasoning skills.

First, the denominator matters. A lottery advertisement may report several winners without stating how many tickets were sold. A health report may state that cases increased without explaining the size of the population at risk.

Second, absolute probability matters. Describing a change as a large relative increase can sound alarming even when the original probability was extremely small. Readers should ask for both relative and absolute measures.

Third, samples and visible cases may be unrepresentative. Lottery winners receive publicity, while losers remain mostly unseen. People who answer health surveys may differ from those who do not.

Fourth, uncertainty cannot be eliminated. A probability does not predict exactly what will happen to one lottery ticket, and a population risk does not determine an individual's future health. Statistics describes patterns and uncertainty rather than guaranteeing specific outcomes.

Finally, communication must be ethical. Lottery organizations should not exaggerate the likelihood of wealth, and health researchers should not overstate observational associations. Accurate numbers

can still mislead when they are presented without context.

Conclusion

The questions “Is your life healthy?” and “Are lotteries fair?” demonstrate how statistical reasoning supports better decisions.

Lottery advertisements may make wealth appear close and attainable, but the probability of winning a major jackpot is extremely small. A lottery can use an honest random drawing and still provide a negative expected financial return. Voluntary participation and public revenue are relevant benefits, but they do not automatically make every aspect of the system fair. A complete analysis must consider informed choice, advertising, distributional effects, problem gambling, and the use of government revenue.

Lotteries generate billions of dollars for states, but it is inaccurate to assume that this revenue necessarily lowers taxes. Budget outcomes depend on government decisions, and lottery funding may supplement or replace other sources without producing a direct tax reduction.

Health studies require equally careful interpretation. Researchers must define health, identify the target population, select an appropriate sample, measure behavior accurately, and distinguish association from causation. A Colorado health survey can provide useful evidence, but its credibility depends on representativeness, response rates, question wording, measurement quality, and appropriate statistical analysis.

Statistics does not provide automatic answers. It supplies tools for evaluating evidence. People who understand probability, sampling, expected value, confounding, effect size, and uncertainty are better prepared to judge advertisements, health claims, government policies, and personal decisions.

References

Bennett, J. O., Briggs, W. L., & Triola, M. F. (2018). *Statistical reasoning for everyday life* (5th ed.). Pearson.

Carruthers, C. K., & Wanamaker, M. H. (2020). Are “education lotteries” less regressive? Evidence from Texas. *Southern Economic Journal*, 86(3), 1019–1040.
<https://doi.org/10.1002/soej.12411>

Kearney, M. S. (2005). State lotteries and consumer behavior. *Journal of Public Economics*, 89(11–12), 2269–2299. <https://doi.org/10.1016/j.jpubeco.2004.09.004>

Lockwood, B. B., Newberry, P., & Taubinsky, D. (2021). *What drives demand for state-run lotteries? Evidence and welfare implications* (NBER Working Paper No. 28975). National Bureau of Economic

Research. <https://doi.org/10.3386/w28975>

Multi-State Lottery Association. (n.d.). *Powerball prize chart*.
<https://www.powerball.com/powerball-prize-chart>

Nurses' Health Study. (n.d.). *History*. <https://nurseshealthstudy.org/about-nhs/history>

Stampfer, M. J., Hu, F. B., Manson, J. E., Rimm, E. B., & Willett, W. C. (2000). Primary prevention of coronary heart disease in women through diet and lifestyle. *The New England Journal of Medicine*, 343(1), 16-22. <https://doi.org/10.1056/NEJM200007063430103>

U.S. Census Bureau. (2026, April 8). *State lottery ticket sales soar as prizes get larger*.
<https://www.census.gov/library/stories/2026/04/state-lottery-ticket-sales-soar.html>