Probability and Statistics (EPRST)

Lecture 7

The expectations of some important distributions

- if $X \sim \mathrm{bin}(n, p)$, then $\mathbb{E}X = ?$,
- if $X \sim \mathrm{geom}(p)$, then $\mathbb{E}X = ?$,
- if $X \sim \mathrm{Poiss}(\lambda)$, then $\mathbb{E}X = ?$,
- if $X \sim \mathrm{U}(a, b)$, then $\mathbb{E}X = ?$,
- if $X \sim \mathrm{Exp}\,(\lambda)$, then $\mathbb{E}X = ?$,
- if $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then $\mathbb{E}X = ?$

Sample mean

In statistics, a central problem is how to use data to estimate unknown parameters of a distribution. It is especially common to want to estimate the mean of a distribution (=the expected value of a random variable).

If the data are n values of a random variable X (generated independently), then the most natural way to estimate $\mathbb{E}X$ is simply to average the values, taking the arithmetic mean. For example, if the observed data are $3, 1, 1, 5$, then a simple, natural way to estimate the mean of the distribution that generated the data is to use

$$\frac{3+1+1+5}{4} = 2.5.$$

This is called the **sample mean**. The sample mean is an empirical estimate of the theoretical expected value (sometimes called the **population mean** or **true mean**) .

Sample mean - cont'd

So if $x_1, \ldots, x_n$ are some (random) values of a random variable X (independently generated) then the sample mean is

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Example 1

We rolled a symmetric die 9 times and got: 3, 1, 2, 5, 1, 4, 3, 2, 6. The sample mean, computed from this particular sample is

$$\bar{X}_9 = \frac{1}{9}(3 + 1 + 2 + 5 + 1 + 4 + 3 + 2 + 6) = 3.$$

The population mean $\mathbb{E}X = 3.5$.

The expectation of a function of random variable

If $Y = g(X)$ then

$$\mathbb{E}(Y) = \mathbb{E}g(X) =$$
$$= \begin{cases} \sum_{x_i \in S} g(x_i) \cdot \mathbb{P}(X = x_i), & \text{discrete case,} \\ \int_{\mathbb{R}} g(x) \cdot f(x)\mathrm{d}x, & \text{continuous case,} \end{cases}$$

(if the series or the integral above are absolutely convergent).
As previously, if $g(X) \geq 0$ (that is $\mathbb{P}(g(X) \geq 0) = 1$), then if the series (or the integral) diverge to ∞ - then we define $\mathbb{E}g(X) = \infty$.

Some examples

Example

Suppose $\mathbb{P}(X = -1) = \mathbb{P}(X = 1) = 1/4$, $\mathbb{P}(X = 0) = 1/2$. *What is* $\mathbb{E}X^2$*? Does it equal* $(\mathbb{E}X)^2$*?*

Example

Compute $\mathbb{E}X^2$ *for* $X \sim U(0, 1)$. *Is it the same as* $(\mathbb{E}X)^2$*?*

The expectation of a function of random variable - some special cases

Definition

- *If $g(x) = x^k$ for some k, then*

$$\mathbb{E}g(X) = \mathbb{E}X^k$$

  *is called the k-th **moment** of X.*
- *If $g(x) = (x - \mathbb{E}X)^2$, then*

$$\mathbb{E}g(X) = \mathbb{E}(X - \mathbb{E}X)^2$$

  *is called the **variance** of X (denoted $\operatorname{Var} X$).*
- $\operatorname{D} X = \sqrt{\operatorname{Var} X}$ *is the **standard deviation** of X.*

The variance - some properties

So the variance of X is defined as

$$\operatorname{Var} X = \mathbb{E}(X - \mathbb{E}X)^2.$$

In order for $\operatorname{Var} X$ to exist, the second moment of X must be finite: $\mathbb{E}X^2 < \infty$.
Properties of the variance:

- $\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2$,
- $\operatorname{Var} X \geq 0$,
- $\operatorname{Var} X = 0$ iff X has a one-point distribution,
- $\operatorname{Var}(X + b) = \operatorname{Var} X$ for every number b,
- $\operatorname{Var}(aX) = a^2 \operatorname{Var} X$ for every number a.

In particular, for all $a, b \in \mathbb{R}$

$$\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X).$$

Computing the variance - some examples

- If $\mathbb{P}(X = 1) = p = 1 - \mathbb{P}(X = 0)$, then $\operatorname{Var} X =?$
- If $X \sim \mathrm{U}(0, 1)$, then $\operatorname{Var} X =?$
- If $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then $\operatorname{Var} X =?$

Computing the variance - some examples

Sort the variances of X, Y, Z and W in the increasing order, if

- X, with the values: $\{1, 2, 3, 4, 5\}$

$$\mathbb{P}(X = 1) = \ldots = \mathbb{P}(X = 5) = \frac{1}{5},$$

- Y, the values $\{1, 2, 3, 4, 5\}$:

$$\mathbb{P}(Y = 1) = \mathbb{P}(Y = 5) = \frac{1}{10}, \ \mathbb{P}(Y = 2) = \mathbb{P}(Y = 4) = \frac{2}{10},$$
$$\mathbb{P}(Y = 3) = \frac{4}{10}$$

- Z, the values $\{1, 5\}$:

$$\mathbb{P}(Z = 1) = \mathbb{P}(Z = 5) = \frac{1}{2},$$

- W, the only value $\{3\}$:

$$\mathbb{P}(W = 3) = 1.$$

Variances of the important named distributions

- if $X \sim \mathrm{bin}(n, p)$, then $\mathrm{Var}\, X = np(1-p)$
- if $X \sim \mathrm{geom}(p)$, then $\mathrm{Var}\, X = (1-p)/p^2$
- if $X \sim \mathrm{Poiss}(\lambda)$, then $\mathrm{Var}\, X = \lambda$
- if $X \sim \mathrm{U}(a, b)$, then $\mathrm{Var}\, X = (b-a)^2/12$
- if $X \sim \mathrm{Exp}\,(\lambda)$, then $\mathrm{Var}\, X = 1/\lambda^2$
- if $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, then $\mathrm{Var}\, X = \sigma^2$

Sample variance and sample standard deviation

If $x_1, \dots, x_n$ are some (random) values of a random variable X (independently generated) then a natural estimate of $\mathbb{E}g(X)$ (g is any function) is the arithmetic mean of the values $g(x_1), \dots, g(x_n)$:

$$\frac{1}{n}\sum_{i=1}^{n} g(x_i).$$

However, the **sample variance** is defined as

$$S_n^2 := \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{X}_n\right)^2$$

The **sample standard deviation** is the square root of the sample variance.

Sample variance - cont'd

The idea of the definition is to mimic the formula

$$\operatorname{Var} X = \mathbb{E}\,(X - \mathbb{E}X)^2$$

by averaging the squared distances of the x_i from the sample mean, except with $n-1$ rather than n in the denominator.
The motivation for the $n-1$ is that this makes S_n^2 **unbiased** for estimating $\operatorname{Var} X$, that is it is correct on average.

Quantiles

Definition

*Let $q \in (0,1)$ and X - a random variable. Number a_q is q-**quantile** of the distribution of X, if*

$$\mathbb{P}(X \leq a_q) \geq q \quad and \quad \mathbb{P}(X \geq a_q) \geq 1 - q,$$

or, equivalently,

$$\mathbb{P}(X < a_q) \leq q \leq \mathbb{P}(X \leq a_q).$$

- *For $q = 1/2$, q-quantile is called **median** (denoted: $\operatorname{med} X$).*
- ***Quartiles** q-quantiles with $q = 1/4, 1/2, 3/4$.*
- *Also **deciles** and **percentiles** are frequently considered.*

Quantiles - some examples

Example

- *if* $\mathbb{P}(X = -1) = \mathbb{P}(X = 1) = 1/4$, $\mathbb{P}(X = 0) = 1/2$, *then* $\operatorname{med} X = ?$
- *if* X *has a discrete uniform distribution on the set* $\{1, 2, 3, 4\}$, *then* $\operatorname{med} X = ?$

Quantiles - cont'd

If the cumulative distribution function of the distribution of a random variable X is a function which is continuous and strictly increasing on an interval (a, b), with $-\infty \leq a < b \leq \infty$, then the definition of quantile becomes simpler - a number c is a q-quantile, if

$$\mathbb{P}(X \leq c) = F_X(c) = q,$$

so

$$c = F_X^{-1}(q).$$

Example

- *If* $X \sim \mathcal{N}(0, 1)$, *then* $\operatorname{med} X =?$
- *If* X *has a Cauchy distribution, then* $\operatorname{med} X =?$

Median vs expectation

The following assertions hold:

- if X is a random variable such that $\mathbb{E}|X| < \infty$, then function

$$f_1(a) = \mathbb{E}|X - a|$$

  attains its minimal value at

$$a = \operatorname{med} X.$$

- if X is a random variable such that $\mathbb{E}X^2 < \infty$, then function

$$f_2(a) := \mathbb{E}(X - a)^2$$

  attains its minimal value at

$$a = \mathbb{E}X.$$