

Lecture 18: Oct 3

Last time

- Practice examples

Today

- Midterm 1 practice exam posted on canvas
- Moments and moment generating function

Moments

Example (Minimizing distance) The expected value of a random variable has another property, one that we can think of as relating to the interpretation of EX as a good guess at a value of X .

Suppose we measure the distance between a random variable X and a constant b by $(X - b)^2$. The closer b is to X , the smaller this quantity is. We can now determine the value of b that minimizes $E[(X - b)^2]$ and, hence, will provide us with a good predictor of X . (Note that it does no good to look for a value of b that minimizes $(X - b)^2$, since the answer would depend on X , making it a useless predictor of X .)

We could proceed with the minimization of $E(X - b)^2$ by using calculus, but there is a simpler method:

$$\begin{aligned} E(X - b)^2 &= E(X - EX + EX - b)^2 \\ &= E[(X - EX) + (EX - b)]^2 \\ &= E(X - EX)^2 + (EX - b)^2 + 2E[(X - EX)(EX - b)], \end{aligned}$$

where we have expanded the square. Note that $E[(X - EX)(EX - b)] = (EX - b)E(X - EX) = 0$, since $EX - b$ is constant and comes out of the expectation, $E(X - EX) = EX - EX = 0$. This means

$$E(X - b)^2 = E(X - EX)^2 + (EX - b)^2.$$

Such that $E(X - b)^2$ is minimized at $b = EX$. And $E(X - EX)^2$ is actually the variance of X ($Var X = E(X - EX)^2$).

The various moments of a distribution are an important class of expectations.

Definition For each integer n , the n th *moment* of X (or $F_X(x)$), μ'_n , is

$$\mu'_n = EX^n.$$

The n th *central moment* of X , μ_n , is

$$\mu_n = E(X - \mu)^n,$$

where $\mu = \mu'_1 = EX$.

Notes:

- $\mu'_0 = EX^0 = 1$
- μ'_1 is the *mean*, usually denoted by μ .
- $\mu_0 = E(X - \mu)^0 = 1$
- $\mu_1 = 0$
- $\mu_2 = E(X - EX)^2$ is the *variance*
- $\mu_3 = E(X - EX)^3$ is related to the *skewness*.
- $\mu_4 = E(X - EX)^4$ is related to the *kurtosis*.

Definition The *variance* of a random variable X is its second central moment, $\text{Var}(X) = E[(X - EX)^2]$. The positive square root of $\text{Var}(X)$ is the *standard deviation* of X .

The variance gives a measure of the degree of spread of a distribution around its mean. Figure 18.1 shows a plot of two samples, one sample draws 100 numbers from a normal distribution with mean 0 and variance 1, $N(0, 1)$. The other sample draws 100 numbers from a normal distribution with mean 0 and variance 100, $N(0, 100)$.

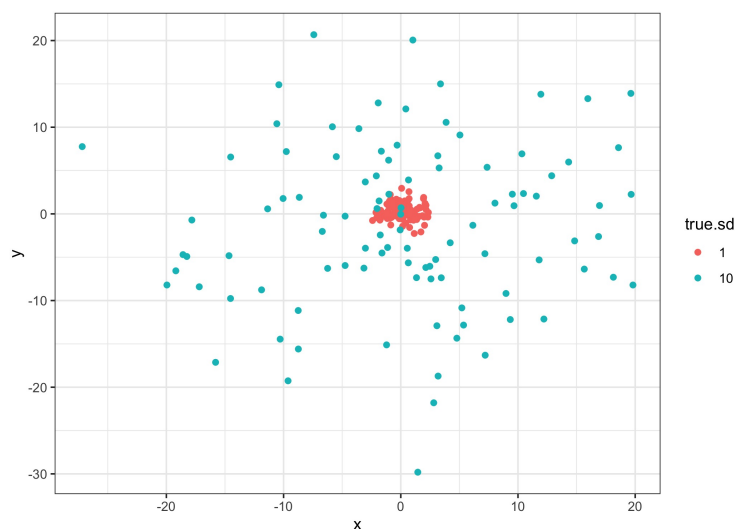


Figure 18.1: Figure 2.1.2. Two samples of 100 numbers drawn from $N(0, 1)$ and $N(0, 100)$.

Example (Exponential variance) Let X have the exponential(λ) distribution. We can calculate the variance of X now.

Solution:

$$\begin{aligned}\text{Var}(X) &= E(X - \lambda)^2 \\ &= \int_0^\infty (x - \lambda)^2 \frac{1}{\lambda} e^{-x/\lambda} dx\end{aligned}$$

$$\begin{aligned}\text{Var}(X) &= \int_0^\infty (x^2 - 2x\lambda + \lambda^2) \frac{1}{\lambda} e^{-x/\lambda} dx \\ &= \int_0^\infty x^2 \frac{1}{\lambda} e^{-x/\lambda} dx - 2 \int_0^\infty x\lambda \frac{1}{\lambda} e^{-x/\lambda} dx + \lambda^2 \\ &= EX^2 - \lambda^2 \\ &= \lambda^2\end{aligned}$$

Theorem If X is a random variable with finite variance, then for any constants a and b ,

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

Proof:

From the definition, we have

$$\begin{aligned}\text{Var}(aX + b) &= E[(aX + b) - E(aX + b)]^2 \\ &= E(aX - aEX)^2 \\ &= a^2 E(X - EX)^2 \\ &= a^2 \text{Var}(X).\end{aligned}$$

It is sometimes to use an alternative formula for the variance, given by

$$\text{Var}(X) = E(X^2) - (EX)^2,$$

which is easily established by

$$\begin{aligned}\text{Var}(X) &= E(X - EX)^2 = E[X^2 - 2XEX + (EX)^2] \\ &= EX^2 - 2(EX)^2 + (EX)^2 \\ &= EX^2 - (EX)^2.\end{aligned}$$

Example (Binomial variance) Let $X \sim \text{Binomial}(n, p)$, that is ,

$$\Pr(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}.$$

What is the variance of X ?

Solutions:

Method #1:

We want to find EX^2 first. We use the

$$EX^2 = \sum_{x=0}^n x^2 \binom{n}{x} p^x (1 - p)^{n-x}.$$

we use the same property $x^2 \binom{n}{x} = xn \binom{n-1}{x-1}$. We then have

$$\begin{aligned}
EX^2 &= n \sum_{x=1}^n x \binom{n-1}{x-1} p^x (1-p)^{n-x} \\
&= n \sum_{y=0}^{n-1} (y+1) \binom{n-1}{y} p^{y+1} (1-p)^{n-1-y} \\
&= np \sum_{y=0}^{n-1} y \binom{n-1}{y} p^y (1-p)^{n-1-y} + np \sum_{y=0}^{n-1} \binom{n-1}{y} p^y (1-p)^{n-1-y} \\
&= np \cdot (n-1)p + np \\
&= n(n-1)p^2 + np.
\end{aligned}$$

And now

$$\begin{aligned}
\text{Var}(X) &= EX^2 - (EX)^2 \\
&= n(n-1)p^2 + np - (np)^2 \\
&= np - np^2 \\
&= np(1-p).
\end{aligned}$$

Method #2:

Recall that we could write $X = \sum_{i=1}^n I_i$, where $I_i \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$. Then

$$\begin{aligned}
\text{Var}(X) &= \text{Var}\left(\sum_{i=1}^n I_i\right) \\
&= \sum_{i=1}^n \text{Var}(I_i) \quad (I_i\text{'s are independent}) \\
&= n \text{Var}(I_1) \quad (I_i\text{'s are identically distributed}) \\
&= n [E(I_1^2) - (EI_1)^2] \\
&= n [p - p^2] \\
&= np(1-p).
\end{aligned}$$

Definition Let X be a random variable with cdf F_X . The *moment generating function (mgf)* of X (or F_X), denoted by $M_X(t)$, is

$$M_X(t) = Ee^{tX},$$

provided that the expectation exists for t in some neighborhood of 0. That is, there is an $h > 0$ such that, for all t in $-h < t < h$, Ee^{tX} exists. If the expectation does not exist in a neighborhood of 0, we say that the moment generating function does not exist.

More explicitly, we can write the mgf of X as

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx, \quad \text{if } X \text{ is continuous,}$$

or

$$M_X(t) = \sum_x e^{tx} \Pr(X = x), \quad \text{if } X \text{ is discrete.}$$

It is easy to see how the mgf generates moments as in the following theorem.

Theorem If X has mgf $M_X(t)$, then

$$EX^n = M_X^{(n)}(0),$$

where we define

$$M_X^{(0)} = \frac{d^n}{dt^n} M_X(t) \Big|_{t=0}.$$

That is, the n^{th} moment is equal to the n^{th} derivative of $M_X(t)$ evaluated at $t = 0$.

Proof:

$$\begin{aligned} \frac{d}{dt} M_X(t) &= \frac{d}{dt} \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \\ &= \int_{-\infty}^{\infty} \left(\frac{d}{dt} e^{tx} \right) f_X(x) dx \\ &= \int_{-\infty}^{\infty} (x e^{tx}) f_X(x) dx \\ &= E(X e^{tX}). \end{aligned}$$

Therefore,

$$\frac{d}{dt} M_X(t) \Big|_{t=0} = E(X e^{tX}) \Big|_{t=0} = EX.$$

Proceeding in an analogous manner, we can establish that

$$\frac{d^n}{dt^n} M_X(t) \Big|_{t=0} = E(X^n e^{tX}) \Big|_{t=0} = EX^n.$$