Lecture 19: Oct 11

Last time

- Distribute Midterm Exam 1

Today

- Vote for alternative weighting (remind me 10 min before class ends):
  - Old one still works
  - New: if Midterm exam 2 score is higher than Midterm exam 1 score, then use Midterm exam 2 score for both
  - Final grade will use the higher one
- Moments
- Moment Generating Function

Moments

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

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

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

The nth *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, $\mathrm{Var}\,(X) = E\left[(X - EX)^2\right]$. The positive square root of $\mathrm{Var}\,(X)$ 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, $N(0, 100)$.

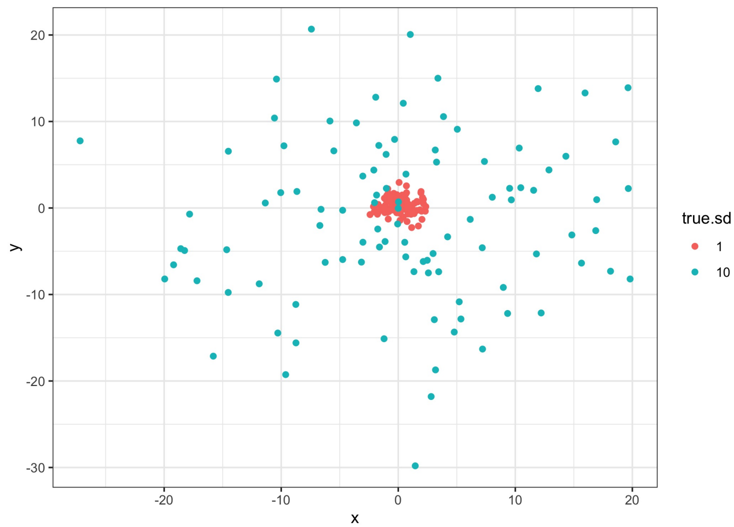

Figure 18.1: Figure 2.1.2. Two samples of 100 numbers drawn from $N(0, 1)$ and $N(0, 100)$.

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

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

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

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

Proof:
From the definition, we have

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

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

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

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

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

What is the variance of X?
Solutions:

Moment Generating Function

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

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

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

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

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

Proof:

Example (Binomial mgf) Let $X \sim Binomial(n,p)$, then its mgf is

$$\begin{aligned} M_X(t) &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x} \\ &= [pe^t + (1-p)]^n. \end{aligned}$$

Theorem Let $F_X(x)$ and $F_Y(y)$ be two cdfs all of whose moments exist.

1. If X and Y have **bounded support**, then $F_X(u) = F_Y(u)$ for all u if and only if $EX^r = EY^r$ for all integers $r = 0, 1, 2, \ldots$.
2. If the moment generating functions exist and $M_X(t) = M_Y(t)$ for all t in some neighborhood of 0, then $F_X(u) = F_Y(u)$ for all u.

Theorem (Convergence of mgfs) Suppose $\{X_i, i = 1, 2, \ldots\}$ is a sequence of random variables, each with mgf $M_{X_i}(t)$. Furthermore, suppose that

$$\lim_{i \to \infty} M_{X_i}(t) = M_X(t), \qquad \text{for all } t \text{ in a neighborhood of } 0,$$

and $M_X(t)$ is an mgf. Then there is a unique cdf F_X whose moments are determined by $M_X(t)$ and, for all x where $F_X(x)$ is continuous, we have

$$\lim_{i \to \infty} F_{X_i}(x) = F_X(x).$$

That is, *convergence*, for $|t| < h$, of mgfs to an mgf implies *convergence* of cdfs.

Poisson approximation One approximation that is usually taught in elementary statistics courses is that binomial probabilities can be approximated by Poisson probabilities. It is taught that the Poisson approximation is valid "when n is large and np is small", and rules of thumb are sometimes given.

The $Poisson(\lambda)$ pmf is given by

$$\Pr(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots,$$

where λ is a positive constant. The approximation states that if $X \sim Binomial(n,p)$ and $Y \sim Poisson(\lambda)$, with $\lambda = np$, then

$$\Pr(X = x) \approx \Pr(Y = x)$$

for large n and small np. We now show that the mgf converge, lending credence to this approximation. Recall that

$$M_X(t) = [pe^t + (1-p)]^n.$$

For the $Poisson(\lambda)$ distribution, we can calculate (HW4, exercise 2.33)

$$M_Y(t) = e^{\lambda(e^t - 1)},$$

and if we define $p = \lambda/n$, then $M_X(t) = [1 + (e^t - 1)\lambda/n]^n$ such that $M_X(t) \to M_Y(t)$ as $n \to \infty$.

Theorem For any constant a and b, the mgf of the random variable $aX + b$ is given by

$$M_{aX+b} = e^{bt} M_X(at).$$

Proof: