6.3: The Mean, Variance and MGF for Several Variables

Taylor

University of Virginia

Definitions

Linear Combo

Let $X_1, \ldots, X_n$ be some r.v.s. Let $a_1, \ldots a_n$ be some real number constants. Then the new random variable

$$Y = a_1 X_1 + \cdots + a_n X_n = \sum_{i=1}^{n} a_i X_i$$

is called a **linear combination**

Note

- $\bar{X}$ is the case when $a_1 = \cdots = a_n = \frac{1}{n}$ (check)
- T_0 is the case when $a_1 = \cdots = a_n = 1$ (check)
- we didn't require all the r.v.s to be independent or identically distributed here

Propositions

linearity of $E(\cdot)$

Assuming all the expectations written down exist

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i)$$

Proof: just properties of integrals/sums

Variance

How $V(\cdot)$ works

Assuming all the expectations and variances written down exist, and assuming all X_i are pairwise independent

$$V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 V(X_i)$$

...and if we can't assume independence...

How $V(\cdot)$ works

Assuming all the expectations and variances written down exist,

$$V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j)$$

Proof: see page 310

Note

- the second one is a more general form than the first (check)
- $Cov(X_i, X_i) = V(X_i)$
- personally I think of $\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j)$ as a quadratic form (but matrix stuff isn't really part of the curriculum)

Example

What if $Y = \sum_{i=1}^{2} a_i X_i = (1)X_1 + (-1)X_2 = X_1 - X_2$ and X_1 is independent from X_2?

Then

$$E(Y) = E(X_1) - E(X_2)$$

and

$$V(X_1 - X_2) = V(X_1) + V(X_2)$$

Theorem

When we take a linear combination of random variables (possibly dependent ones), we have the tools to calculate the mean and variance. We would like to know more, though. We would like to know the entire probability distribution. In the case of normals, this is particularly easy:

The Reproductive Property of Normal r.v.s

If $X_1, \ldots, X_n$ are independent (not necessarily identical) normal rvs, then any linear combination $Y = \sum a_i X_i$ is also normally distributed.

Notes

- Proof comes later after we talk about MGFs
- Actually any joint normal vector stays normal under any linear (affine really) transformation
- We really only care about when we start off with i.i.d. normal rvs in this class, though

MGFs

You learned about these already, but here's the definition again.

$$M_X(t) = E(\exp(tX))$$

Remember you can also identify rvs by their MGF (instead of their density). Sometimes it's easier to work with the MGF.

MGFs

In particular, they help out a lot with linear combinations of *independent* rvs. Let $Y = \sum_i a_i X_i$

$$
\begin{aligned}
E[\exp(tY)] &= E[\exp(t\left\{\sum_i a_i X_i\right\})] \\
&= E[\exp(ta_1 X_1) \cdots \exp(ta_n X_n)] \\
&= E[\exp[(ta_1)(X_1)]] \cdots E[\exp[(ta_n)(X_n)]] \\
&= M_{X_1}(ta_1) \cdots M_{X_n}(ta_n)
\end{aligned}
$$

Example

Let's consider the sum of two independent Poisson count random variables. Let's call them X and Y with parameters λ and ν respectively. What's the distribution of $X + Y$?
FYI the MGF for a Poisson rv with a parameter θ is $\exp[\theta(e^t - 1)]$

Example

$$\begin{aligned} M_{X+Y}(t) &= M_X(t)M_Y(t) \\ &= \exp[\lambda(e^t - 1)]\exp[\nu(e^t - 1)] \\ &= \exp[(e^t - 1)(\nu + \lambda)] \end{aligned}$$

... so it's still a poisson but with parameter $\lambda + \nu$

Proof

Recall

$$M_{\sum_i a_i X_i}(t) = M_{X_1}(ta_1) \cdots M_{X_n}(ta_n).$$

Let's use this to prove the reproductive property of normals. We just need to know what an MGF for a normal distribution looks like. If $X \sim \text{Normal}(\mu, \sigma^2)$, then $M_X(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$

Proof

Let $X_1, \ldots, X_n$ all be independent but follow possibly different normal distributions. Define Y as the linear combination: $Y = \sum_i a_i X_i$. Then

$$
\begin{aligned}
M_{\sum_i a_i X_i}(t) &= \prod_i M_{X_i}(a_i t) \\
&= \prod_i \exp(\mu_i a_i t + \frac{1}{2}\sigma_i^2 (a_i t)^2) \\
&= \exp(\mu^* t + \frac{1}{2}\sigma^{*2} t^2)
\end{aligned}
$$

So Y follows a normal distribution again with mean $\mu^* = \sum_i a_i \mu_i$ and variance $\sigma^{*2} = \sum_i a_i^2 \sigma_i^2$