MA 6.101
Probability and Statistics

Tejas Bodas

Assistant Professor, IIIT Hyderabad

Topics

We have seen

- Conditioning
- Law of iterated Expectations
- Sums of random variables & Convolutions
- Bayes Rule revisited

This class ..

- Some more properties of $E[E[X|Y]]$
- Variance of sums of random variables
- Moment Generating functions

Law of Iterated Expectation revisited

- Recall $E[X] = E[E[X|Y]]$. What are the two expectations w.r.t ?
- Let $g(Y) = E[X|Y]$. Then

$$g(y_1) = E[X|Y = y_1] = \int_x x f_{X|Y}(x|y_1) dx$$

. So the inner expectation is w.r.t X.

- $E[X] = E[g(Y)] = \int_y g(y) f_Y(y) dy$. So the outer expectation is w.r.t Y.

$$\boxed{E[X] = E_Y\left[E_X[X|Y]\right]}$$

Law of Iterated Expectation revisited

$$E[X] = E_Y\left[E_X[X|Y]\right]$$

- What is $E[Xg(Y)|Y]$?
- Note that $E[Xg(Y)|Y = y_1] = g(y_1)E[X|Y = y_1]$.
- Therefore $E[Xg(Y)|Y] = g(Y)E[X|Y]$. In general, we have the following pull through property

$$E[h(X)g(Y)|Y] = g(Y)E[h(X)|Y].$$

Law of Iterated Expectation revisited

$$E[X] = E_Y\left[E_X[X|Y]\right]$$

- If X and Y are independent, what is $E[X|Y]$?
- Since X and Y are independent, $E[X|Y = y] = E[X]$ for all y.
- This means $g(Y) = E[X|Y]$ always takes the value of $E[X]$.

In fact, when X and Y are independent, we have

$$E[g(X)|Y] = E[g(X)].$$

Variance of sum of random variables

- Let $X_1, X_2, \ldots X_n$ be possibly dependent and non-identical random variables.
- Lets say you know the joint pdf/pmf for every pair of random variables from this collection.
- AIM: Calculate $Var(Z)$ where $Z = \sum_{i=1}^{n} a_i X_i$ for some scalars a_i.

Variance of sum of random variables

- Recall $Var(X) = E[X - E[X]]^2 = E[X^2] - E[X]^2$.
- Also recall $Cov(X, Y) = E[XY] - E[X]E[Y]$.
- Following properties of covariance follow (HW)

1. $Cov(X, X) = Var(X)$
2. If X, Y are independent, $Cov(X, Y) = 0$.
3. $Cov(X, Y) = Cov(Y, X)$
4. $Cov(aX, Y) = aCov(X, Y)$
5. $Cov(X + a, Y) = Cov(X, Y)$
6. $Cov(X + Z, Y) = Cov(X, Y) + Cov(Z, Y)$
7. $Cov\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i b_j Cov(X_i, Y_j)$

Variance of sum of random variables

- AIM: Calculate $Var(Z)$ where $Z = \sum_{i=1}^{n} a_i X_i$ for some scalars a_i.
- $Var(Z) = Cov(Z, Z)$ and therefore

$$
\begin{aligned}
Cov\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} a_j X_j\right) &= \sum_{i=1}^{m} \sum_{j=1}^{n} a_i a_j Cov(X_i, X_j) \\
&= \sum_{i=1}^{n} a_i^2 Var(X_i) \\
&+ \sum_{(i,j):i \neq j} a_i a_j Cov(X_i, X_j)
\end{aligned}
$$

$$
\boxed{Var(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum_{(i,j):i \neq j} a_i a_j Cov(X_i, X_j)}
$$

Variance of sum of random variables

$$Var(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum_{(i,j):i \neq j} a_i a_j Cov(X_i, X_j)$$

- Show that $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
- Now if $X_i's$ are independent, what is $Var(Z)$?
- Let $\{X_i, i = 1, 2, \ldots n\}$ be i.i.d and consider $S_n = \frac{\sum_{i=1}^{n} X_i}{n}$.
- Show that $Var(S_n) = \frac{Var(X)}{n}$

Moment generating function

- The moment generating function (MGF) of a random variable X is a function $M_X : \mathbb{R} \to [0, \infty]$ defined by $M_X(t) = E[e^{tX}]$.
- If X is discrete, $M_X(t) = \sum_{x \in \Omega'} e^{tx} p_X(x)$.
- If X is continuous, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx$.
- Define $D_X := \{t : M_X(t) < \infty\}$. D_X is called the region of convergence (ROC). $t = 0$ is always part of ROC.
- Find MGF of Z where Z is a Bernoulli(p) random variable.

MGF examples

- For $Exp(\lambda)$ variable, $M_X(t) = \frac{\lambda}{\lambda - t}$ for $\lambda < t$.
- For $Z \sim \mathcal{N}(\mu, \sigma^2)$, we have $M_Z(t) = e^{(\mu t + \frac{1}{2}\sigma^2 t^2)}$
- `https://proofwiki.org/wiki/Moment_Generating_Function_of_Gaussian_Distribution`
- HW: Find the MGF for a random variable X that has the following distributions: Binomial(n,p), Normal $\mathcal{N}(0, 1)$, Poisson(λ)

MGF

- If $M_X(t)$ is finite for all $|t| \leq \epsilon$ and for some $\epsilon > 0$ then $M_X(t)$ is infinitely differentiable on $(-\epsilon, \epsilon)$. (Property without proof)
- Let ${M_X}^{(r)}(t) := \frac{d^r}{dt^r} M_X(t)$ (r^{th}-derivative of $M_X(t)$)
- Intuitively, one can see that ${M_X}^{(r)}(t) = E[e^{tX} X^r]$ for all r.
- $E[X^r] = {M_X}^{(r)}(0)$
- HW: Work out these things for $Exp(\lambda)$
- HW: Find MGF for all random variables studied till now

MGF of Sums of independent random variable

- Consider $Z = X + Y$. What is the pdf of Z when X and Y?
- Let $M_X(t)$ and $M_Y(t)$ be their MGF's. What is $M_Z(t)$?
- $M_Z(t) = E[e^{Zt}] = E[e^{(X+Y)t}]$.
- $M_Z(t) = E[e^{Xt}.e^{Yt}]$.
- If X and Y are independent, $E[XY] = E[X]E[Y]$ and $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$.
- $M_Z(t) = E[e^{Xt}].E[e^{Yt}]$.

$$\boxed{M_Z(t) = M_X(t)M_Y(t).}$$

MGF of Sums of independent random variable

- Consider $Z = X + Y$. What is the MGF of Z when X and Y?

$$\boxed{M_Z(t) = M_X(t)M_Y(t).}$$

- What about $M_Z(t)$ when $Z = X_1 + X_2 + \dots X_n$ and X_i are iid.?
- $M_Z(t) = (M_X(t))^n$.
- What about $M_Z(t)$ when $Z = X_1 + X_2 + \dots X_N$ where N is a positive discrete random variable?
- $M_Z(t) = E[e^{tZ}] = E_N[E[e^{tZ}|N]] = E_N((M_X(t))^N)$.
- $M_Z(t) = \sum_n p_N(n) M_X(t)^n$
- HW: Prove that $M_Z(t) = M_N(log M_X(t))$