Section 3: Expectation

ARE 210

September 12, 2017

- Introduction (10 min)
- Quick tips (5 min)
- Practice questions (35 min)

The section notes are available on the section Github at github.com/johnloeser/are210 in the "section3" folder.

1 Definitions

- Let μ be a measure over a measurable space (Ω, F)
  - $\int \mathbf{1}_A(\omega)\mu(d\omega) = \mu(A)$, where $\mathbf{1}_A(\omega) = \mathbf{1}\{\omega \in A\}$
  - Let $f(\omega) = \sum_j a_j \mathbf{1}_{A_j}(\omega)$, then $\int f(\omega)\mu(d\omega) = \sum_j a_j \mu(A_j)$. We call functions of the form of f **step functions**
  - Let $f_n \to f$ be a sequence of step functions converging in the limit to f, a measurable function. $\int f(\omega)\mu(d\omega) \equiv \lim_{n\to\infty} \int f_n(\omega)\mu(d\omega)$
  - $\int_A f(\omega)\mu(d\omega) \equiv \int f(\omega)\mathbf{1}_A(\omega)\mu(d\omega)$
  - From here on out, we'll almost always use Lebesgue measure over the Borel σ-algebra, which means Lebesgue integration is equivalent to traditional Riemann integration, and standard results hold (Fundamental Theorem of Calculus, Change of Variable, ...)
    - Let X be a random variable with probability measure P_X and density f, and g a function. Then $\int g(x)P_X(dx) = \int g(x)f(x)dx$
    - Need to be careful when working with discrete or mixed random variables
    - Need to be careful when working with infinities - break integral into set when f is positive and set when f is negative, if both integrals infinite then integral undefined
- A holds almost surely $\Leftrightarrow P(A) = 1$
- $\mathbf{E}[X] \equiv \int_\Omega X(\omega)P[d\omega]$
  - $\mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y]$
  - $X \sim F \Rightarrow \mathbf{E}[X] = \int_0^\infty [1 - F(t)]dt - \int_{-\infty}^0 F(t)dt$
  - $X \perp Y \Rightarrow \mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$

- nth moment of $X \equiv \mathbf{E}[X^n]$
- nth centered moment of $X \equiv \mathbf{E}[(X - \mathbf{E}[X])^n]$
- $\text{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])^T]$
  - Can verify Cov is linear from $\mathbf{E}$ properties
- $M_X(t) = \mathbf{E}[\exp(tX)]$ is the **moment generating function**
  - If the MGF is defined in a neighborhood of 0 for X and Y, then X and Y have the same distribution if and only if M_X and M_Y are identical in a neighborhood of 0
- $\Psi_X(t) = \mathbf{E}[\exp(itX)]$ is the **characteristic function**
  - X and Y are identically distributed if and only if Ψ_X and Ψ_Y are identical
  - $\Psi_{aX+b}(t) = \exp(itb)\Psi_X(at)$
  - $X \perp Y \Rightarrow \Psi_{X+Y}(t) = \Psi_X(t)\Psi_Y(t)$
- **Markov's Inequality**: $P[|X| > b] \leq \frac{\mathbf{E}[|X|]}{b}$
- **Chebyshev's Inequality**: $P[|X - \mathbf{E}[X]| > \epsilon] \leq \frac{\mathbf{V}[X]}{\epsilon^2}$
- **Jensen's Inequality**: Let g be a convex function, X a random variable, then $\mathbf{E}[g(X)] \geq g(\mathbf{E}[X])$

2 Some helpful tips

- For our purposes, when taking expectations, we'll almost always be doing so with respect to a random variable with a continuous, discrete, or mixed distribution over $\mathbf{R}$. We'll see next week that we can reduce the mixed case to the discrete and continuous case
  - Expectation of a discrete RV just involves sums
  - Expectation of a continuous RV can be done using Riemann integration
  - In some cases, we can do either using characteristic functions or MGF/Laplace transform
- (MGF:Characteristic Function)::(Taylor Series:Fourier Series)
- We've now covered material needed to answer 1-9 on PS2. You can try 10-14 however - these questions are on conditional expectation. However, we can think about $Y|X$ as a random variable just like Y - we can calculate its distribution using Bayes' Rule, which means we can calculate its CDF. As a result, you have all the tools you need to answer these questions.

3 Practice questions

1) **Expectation practice**: Let $X \sim N(0,1)$.

a) Calculate $\Psi_X(t)$.

b) Calculate $\mathbf{E}[X^3]$.

2) **Lee bounds**: Let $F(y) = pM(y) + (1-p)N(y)$, for $p \in [0,1]$, and let $G(y) = \max\left\{0, \frac{F(y)-p}{1-p}\right\}$. Show that $\int yG(dy) \geq \int yN(dy)$.

3) **Characteristic function of the sample mean**: Let $\{X_i\}_{i=1}^n \sim$ iid with mean μ and variance σ^2. Let $\overline{X}_n \equiv \frac{1}{n}\sum_{i=1}^n X_i$. Let $Y_n \equiv \frac{\overline{X}_n - \mathbf{E}[\overline{X}_n]}{\sqrt{\mathbf{V}[\overline{X}_n]}}$. Derive the characteristic function of Y_n, $\Psi_{Y_n}(t)$.

Hint: Use $\Psi_{Y_n}(t) = \mathbf{E}[\exp(itY_n)]$ and $\Psi_{aX+b}(t) = \exp(itb)\Psi_X(at)$.

4) **Visual intuition of MGF and characteristic function**: Define $M_X(t) = \mathbf{E}[\exp(tX)]$ and $\Psi_X(t) = \mathbf{E}[\exp(itX)]$.

a) Take Taylor series expansions of M_X and Ψ_X around $t = 0$.

b) What's the first derivative of M_X at 0? The second derivative? What's the first derivative of Ψ_X at 0? The second derivative?

c) Graph $\Psi_X(t)$ for a few small t for $X \sim$ Bernoulli(1/3) ($\Psi_X(t) = \frac{1}{3}[\cos t + i \sin t] + \frac{2}{3}$) and for $X \sim N(0,1)$ ($\Psi_X(t) = \exp(-t^2/2)$). Think about this and the fact that $\Psi_X(t) = \mathbf{E}[\cos[tX]] + i\mathbf{E}[\sin[tX]]$.

5) **Chebyshev's Inequality**: Prove that $P[|X - \mathbf{E}[X]| > \epsilon] \leq \frac{\mathbf{V}[X]}{\epsilon^2}$.

Hint: Use the trick from the proof of Markov's Inequality, that $|X| \geq b\mathbf{1}[|X| \geq b]$, and apply it to $Z = (X - \mathbf{E}[X])^2$.