Section 3: Expectation (solutions)

ARE 210

September 12, 2017

1) **Expectation practice**: Let $X \sim N(0,1)$. Calculate $\mathbf{E}[X^3]$.

For the easy way, we can just use the fact that the characteristic function of the standard normal is $\exp(-t^2/2)$. The third derivative of the characteristic function is $-i$ times the third moment. This third derivative is 0, which tells us that $\mathbf{E}[X^3] = 0$.

For the hard way,

$$\begin{aligned}
\mathbf{E}[X^3] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^3 e^{-x^2} dx \\
&= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{0} x^3 e^{-x^2} dx + \int_{0}^{\infty} x^3 e^{-x^2} dx \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(-\int_{0}^{\infty} x^3 e^{-x^2} dx + \int_{0}^{\infty} x^3 e^{-x^2} dx \right)
\end{aligned}$$

Note that the two terms in the parentheses cancel to 0 as long as they are finite. So $\mathbf{E}[X^3]$ is either 0 or undefined. We next calculate one of these terms to verify that its 0.

$$\begin{aligned}
\int_0^{\infty} x^3 e^{-x^2} dx &= -\frac{1}{2} x^2 e^{-x^2} \Big|_0^{\infty} + \int_0^{\infty} x e^{-x^2} dx \\
&= \frac{1}{2} \int_0^{\infty} 2x e^{-x^2} dx \\
&= \frac{1}{2} \int_0^{\infty} e^{-y} dy \\
&= \frac{1}{2}
\end{aligned}$$

These terms are finite, so $\mathbf{E}[X^3] = 0$.

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 $\int yG(dy) \geq \int yN(dy)$.

First, we use our formulas for integration in terms of the CDF.

$$
\begin{aligned}
\int yG(dy) &= \int_0^\infty [1-G(y)]dy - \int_{-\infty}^0 G(y)dy \\
\int yN(dy) &= \int_0^\infty [1-N(y)]dy - \int_{-\infty}^0 N(y)dy \\
\int yG(dy) - \int yN(dy) &= \int_0^\infty [N(y)-G(y)]dy - \int_{-\infty}^0 [G(y)-N(y)]dy \\
&= \int_{-\infty}^\infty [N(y)-G(y)]dy
\end{aligned}
$$

Intuitively, it suffices to show that $N(y) \geq G(y)$. First, note that $F(y) \leq (1-p)N(y) + p$.[1] This implies that $G(y) \leq \max\{0, N(y)\} = N(y)$.

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 $\Psi_{Y_n}(t) = \mathbf{E}[\exp(itY_n)]$, $\Psi_{aX+b}(t) = \exp(itb)\Psi_X(at)$, and $\Psi_{A+B}(t) = \Psi_A(t)\Psi_B(t)$ when A and B are independent.

First, $\mathbf{E}[\overline{X}_n] = \frac{1}{n}\sum_{i=1}^n \mathbf{E}[X_i] = \mu$ and $\mathbf{V}[\overline{X}_n] = \frac{1}{n^2}\sum_{i=1}^n \mathbf{V}[X_i] = \sigma^2/n$.

Using the hint, this gives us

$$
\begin{aligned}
\Psi_{Y_n}(t) &= \exp(-it\sqrt{n}\mu/\sigma)\Psi_{\overline{X}_n}(\sqrt{n}t/\sigma) \\
&= \exp(-it\sqrt{n}\mu/\sigma)\Psi_{\sum_{i=1}^n X_i}(t/(\sqrt{n}\sigma)) \\
&= \exp(-it\sqrt{n}\mu/\sigma)\left(\Psi_{X_i}(t/(\sqrt{n}\sigma))\right)^n
\end{aligned}
$$

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) ($\Psi_X(t) = \frac{1}{3}[\cos t + i\sin t] + \frac{2}{3}$)

[1] To find this, one can look for when a counterexample would exist - G will be largest when M is largest, or when M is always 1. Alternatively, one can think of G as picking the biggest values from the distribution F (from combining N and M), so G will "pick" relatively small values, i.e. G will be large, when M "picks" relatively small values, i.e. M is large.

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

$$M_X(t) = \sum_{j=0}^{\infty} \frac{t^j}{j!}\mathbf{E}[X^j]$$

$$\Psi_X(t) = \sum_{j=0}^{\infty} \frac{(it)^j}{j!}\mathbf{E}[X^j]$$

The first (second) derivative of M_X at 0 is the first (second) moment of X. The first (second) derivative of Ψ_X at 0 is the first (second) moment of X times i (-1).

Basically, we're taking the expectation of the unit circle, where the weights on segments are determined by P_X, but gradually scaling X out using t.

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 $Z = (X - \mathbf{E}[X])^2$.

Let $Z = (X - \mathbf{E}[X])^2$. Then

$$\begin{aligned}
Z &\geq \epsilon^2 \mathbf{1}[Z > \epsilon^2] \\
\mathbf{E}[Z] &\geq \epsilon^2 \mathbf{P}[Z > \epsilon^2] \\
\mathbf{V}[X] &\geq \epsilon^2 \mathbf{P}[Z > \epsilon^2] \\
&= \epsilon^2 \mathbf{P}[|X - \mathbf{E}[X]| > \epsilon] \\
\frac{\mathbf{V}[X]}{\epsilon^2} &\geq \mathbf{P}[|X - \mathbf{E}[X]| > \epsilon]
\end{aligned}$$