The capacity of communication channels has largely been studied in the digital domain, under the premise that sampling, if done above the Nyquist rate of the channel bandwidth, preserves information. However, hardware and power limitations often preclude sampling at this rate, especially for wideband channels. Moreover, the Nyquist rate is not always needed to preserve information; recent results in compressed sensing indicate that when signals exhibit sparse structure they can be sampled below their Nyquist rate without losing information. These ideas give rise to several fundamental questions at the intersection of sampling theory and information theory: how is channel capacity affected by sampling below the channels Nyquist rate; what is the optimal sub-Nyquist rate sampling strategy to maximize capacity; and for what channels is capacity preserved at sub-Nyquist sampling rates. This talk will explore these fundamental questions and provide some answers about fundamental tradeoffs between sampling and capacity.