Lecture 34: Calculus and Statistics

In this lecture, we look at an application of calculus to statistics. We have already defined the probability density function f and its anti-derivative, the cumulative distribution function.
This lecture is given by Brian Lukoff. This document is what I had prepared for this lecture. Brian's handout will be the official course note.

Functions

In statistics, functions appear at many places. First of all for random variables. Then for probability density functions and cumulative probability density functions. In order to compute quantities like expectation and variance, we have to integrate.

The **expectation** of probability density function f is

$$m = \int_{-\infty}^{\infty} x f(x)\, dx \; .$$

The **variance** of probability density function f is

$$\int_{-\infty}^{\infty} (x^2 - m) f(x)\, dx$$

where m is the expectation. The square root of the variance is the **standard deviation**.

The expectation of the normal distribution

$$f(x) = \frac{1}{2\pi\sigma^2} e^{-(x-m)^2/(2\sigma^2)}\, dx$$

is equal m. The standard deviation is σ.

The expectation of the geometric distribution $f(x) = e^{-ax}a$

$$\int x e^{-ax} a\, dx = 1/a \; .$$

The variance of the geometric distribution $f(x) = e^{-ax}a$ is $1/a^2$ and the standard deviation $1/a$.

To see this, compute (remember Tic Tac Toe!)

$$\int x^2 e^{-ax} a\, dx = 2/a^3 \; .$$

x^2	e^{-ax}	
$2x$	$-e^{-ax}/a$	$\oplus$
2	e^{-ax}/a^2	$\ominus$
0	e^{-ax}/a^3	$\oplus$

More Problems

1 The **uniform distribution** on $[a, b]$ is the probability density function which is zero for x outside the interval $[a, b]$ and equal to $1/(b-a)$ for $x \in [a, b]$. Find the mean and standard deviation of this density.

2 The **Laplace distribution** is also called the double exponential distribution. It has the density $1/(2a)e^{-a|x|}$. Find the mean and standard deviation of the Laplace distribution.

3 The **Logarithmic distribution** on $[1, 2]$ has the density $C\log(x)$, where C is a constant which makes it a density. What is the constant C?

4 The **Rayleigh distribution** is the probability density which is 0 for $x < 0$ and $xe^{-x^2/(2s^2)}/s^2$ for $x > 0$). Verify that its mean is $s\sqrt{\pi/2}$ and its variance is $(4-\pi)\sigma^2/2$.

5 The **Maxwell distribution** is the probability density which is zero for $x < 0$ and $\frac{2}{2\pi} x^2 e^{-x^2(/2a^2)}/a^3$ for $x \geq 0$. Verify that its mean is $4a/\sqrt{2\pi}$ and its variance is $a^2(3\pi - 8)/\pi$.