Mathematical Statistics Notes

Zoe Farmer

February 26, 2024

1 Preliminaries

1.1 Expected Value

The expected value for a discrete distribution is defined as

$$E[X] = \sum_x x \cdot P(X = x) = \sum_x x \cdot f(x)$$

For a continuous distribution, the expected value is defined as

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)$$

The following properties hold.

1. Expectation is linear.

$$E[\alpha X + \beta Y] = \alpha E[x] + \beta E[y]$$

2. If X and Y are independent

$$\begin{aligned} E[XY] = &\quad E[X] \cdot E[Y] \\ E[g(X) \cdot h(Y)] = &\quad E[g(X)] \cdot E[h(Y)] \end{aligned}$$

1.2 Variance

If we use μ to denote the mean $E[X]$, then the variance of X is defined by

$$\text{Var}[X] = E[(X - \mu)^2] = E[X^2] - (E[X])^2$$

The following properties hold.

1. $\text{Var}[aX] = a^2 \text{Var}[x]$
2. $\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - E[X]E[Y]$
3. $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y]$

1.3 Indicator Functions

Instead of defining distributions piecewise as we've done in the past we prefer to use indicator functions that take on the values of zero and one.

Let A be a set. The function

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

2 Four Important Tools

2.1 Finding Distributions of Transformations of Random Variables

For discrete distributions it is enough to simply replace the variable with the function. This is most apparent through an example.

$$X \sim bin(n,p) \qquad Y = n - X$$
$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$
$$f_Y(x) = \binom{n}{n-x} p^{n-x} (1-p)^y$$

For continuous distributions it's a little harder.

Let X be a continuous random variable with pdf f_x. Let Y be a random variable defined by $Y = g(X)$ where g is invertible (and differentiable). Then the pdf for Y can be computed as

$$f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot \left| \frac{d}{dy} g^{-1}(y) \right|$$

2.2 Bivariate Transformations

Suppose that X_1 and X_2 are continuous random variables with joint pdf $f_{X_1,X_2}(x_1,x_2)$ and suppose that new random variables Y_1 and Y_2 are defined by

$$Y_1 = g_1(X_1, X_2) \qquad Y_2 = g_2(X_1, X_2)$$

The joint pdf for Y_1 and Y_2 is given by

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}(g_1^{-1}(y_1,y_2), g_2^{-1}(y_1,y_2)) \cdot \left\| \begin{matrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \end{matrix} \right\|$$

If Y_1 is a ratio, it is almost always a good idea to choose Y_2 to be the denominator.

2.3 Order Statistics

We can define a short hand notation for the maximums and minimums.

$$\begin{aligned} X_{(1)} &= \min(X_1, X_2, \dots, X_n) \\ \vdots &= \vdots \\ X_{(n)} &= \max(X_1, X_2, \dots, X_n) \end{aligned}$$

The minimum is defined as

$$f_{X_{(1)}}(x) = n[1 - F(x)]^{n-1} f(x)$$

The maximum is defined as

$$f_{X_{(n)}}(x) = n[F(x)]^{n-1} f(x)$$

2.4 Moment Generating Functions

For a random variable X, the moment generating function denoted by $M(t)$ or $M_X(t)$ is defined as

$$M(t) = E\left[e^{tX}\right]$$

We can use the Law of the Unconscious Statistician.
Let X be a random variable with pdf $f_x(x)$. Let $g(x)$ be some function.
If X is discrete we have

$$E[G(X)] = \sum_x g(x) f_X(x)$$

If X is continuous we have

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

In general, for a random variable X with mgf $M_X(t)$, the kth moment is obtained by $M^{(k)}(0)$, where $M^{(k)}(t)$ is the kth derivative of $M_X(t)$ with respect to t.

The moment generating function for a random variable X uniquely determines its distribution.

If $X_1, \dots, X_n$ are iid random variables from a distribution with moment generating function $M_X(t)$ then the sum $Y = \sum_{i=1}^n X_i$ has moment generating function

$$M_Y(t) = [M_X(t)]^n$$

3 Estimators

4 Distributions

4.1 The Gamma Distribution

4.1.1 The Gamma Function

Defined for $\alpha > 0$ as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \, dx$$

Properties are as follows.

1. $\Gamma(1) = 1$
2. For $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha - 1)$
3. If $n \geq 1$ is an integer, $\Gamma(n) = (n-1)!$

4.2 The Beta Distribution

4.2.1 The Beta Function

The beta function is defined, for $a, b > 0$, as

$$\mathcal{B}(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} \, dx$$

The following property holds.

$$\mathcal{B}(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$