DENSITY NOTATION

JOHN THICKSTUN

1. DENSITIES

Let X be a random variable on $(\Omega, \mathcal{F}, P)$ with pushforward measure X_*P. This measure is called the law of X and

$$P(X \in A) = \int_{\{X \in A\}} dP = \int_A d(X_*P).$$

If the law of X is absolutely continuous with respect to Lebesgue measure, that is $X_*P \ll dx$, then there exists a Radon-Nikodym derivative $d(X_*P)/dx : \mathbb{R} \to \mathbb{R}$ such that

$$\int_A d(X_*P) = \int_A \frac{d(X_*P)}{dx} dx.$$

This derivative is called a density and is commonly denoted by $p(x)$ and we write

$$P(X \in A) = \int_A p(x) dx.$$

Similarly, by the law of the unconscious statistician we write

$$\mathbb{E}X = \int_{-\infty}^{\infty} x p(x) dx.$$

Let $\mathcal{G} \subset \mathcal{F}$ be a (sub)σ-algebra. We can define a unique (up to null sets) conditional expectation $\mathbb{E}\{X|Y\}$, which for any $G \in \mathcal{G}$ satisfies

$$\int_G \mathbb{E}\{X|\mathcal{G}\} dP = \int_G X dP.$$

This relationship is neatly expressed by the following commutative diagram.

$$\begin{tikzcd}
 & \mathbb{R} \\
(\Omega, \mathcal{F}) \arrow[ur, "X"] \arrow[r, "\mathbb{E}\{\cdot|\mathcal{G}\}"'] & (\Omega, \mathcal{G}) \arrow[u, "\mathbb{E}\{X|\mathcal{G}\}"']
\end{tikzcd}$$

That is, $\mathbb{E}\{X|\mathcal{G}\} = Y$ for some $\mathcal{G}$-measurable $Y : \Omega \to \mathbb{R}$. Conversely, if Y is a random variable with $\sigma(Y) = \mathcal{G}$ then we define $\mathbb{E}\{X|Y\} = \mathbb{E}\{X|\mathcal{G}\}$ and for some measurable $h : \mathbb{R} \to \mathbb{R}$, $\mathbb{E}\{X|Y\} = h(Y)$. Furthermore, this allows us to define a pointwise conditional expectation $\mathbb{E}\{X|Y = y\} = h(y)$. We might want to extend this definition into a more general $\mathbb{E}\{X|Y \in A\}$ but here we must be very careful: see, for example, the Borel Kolmogorov paradox.

Recall that $P(X \in A) = \mathbb{E}\{\mathbb{1}_{\{X\in A\}}\}$, and from this observation we can extend our definition of conditional expectation to a definition of conditional probability. In particular, we define $P(\cdot|\mathcal{G}) : \mathcal{F} \times \mathcal{G} \to \mathbb{R}$ by

$$P(X \in A|\mathcal{G}) = \mathbb{E}\{\mathbb{1}_{\{X\in A\}}|\mathcal{G}\}.$$

This construction is not a priori a probability measure. We say that the conditional probability is regular if it in fact does satisfy axioms of probability. We assume from this point on that our conditional probabilities are regular, which allows us to write

$$P(X \in A|\mathcal{G}) = \int_{\{X\in A\}} dP(\cdot|\mathcal{G}) = \int_A dX_*P(\cdot|\mathcal{G}).$$

If $P(\cdot|\mathcal{G}) \ll dx$ then we define a conditional density $p(x|\mathcal{G}) = dX_*P(\cdot|\mathcal{G})/dx$ and write

$$P(X \in A|\mathcal{G}) = \int_A p(x|\mathcal{G})dx$$

If Y generates $\mathcal{G}$ then $\mathbb{E}\{X|\mathcal{G}\} = \mathbb{E}\{X|Y\}$, and we write

$$P(X \in A|Y) = \int_A p(x|y)dx.$$

And likewise,

$$\mathbb{E}\{X|Y\} = \int_{-\infty}^{\infty} xp(x|y)dx.$$

2. MULTIPLE VARIABLES

Now let's consider the case of a pair of random variables $X = (Y, Z)$. Let's assume both Y and Z are defined on $(\Omega, \mathcal{F}, P)$. We can define a joint density in the manner discussed above by

$$P(Y \in A, Z \in B) = \int_{A\times B} p(y, z)d(y, z).$$

From this definition, Fubini's theorem, and the definition of a single-variable density,

$$P(Y \in A) = P(X \in A \times \mathbb{R}) = \int_{A\times\mathbb{R}} p(y, z)d(y, z)$$

$$= \int_A \int_{-\infty}^{\infty} p(y, z)dzdy = \int_A p(y)dy.$$

This holds for all A and it follows that a.s.

$$p(y) = \int_{-\infty}^{\infty} p(y, z)dz, \quad p(z) = \int_{-\infty}^{\infty} p(z, y)dz.$$

These are called the marginals of Y and Z. Note that we have slightly abused notation: $p(y)$ and $p(z)$ are distinct functions, distinguished lexicographically by the name of their input variable. By our definitions of densities and conditional probabilities in section 1,

$$P(Y \in A, Z \in B) = \int_B P(Y \in A|Z) d(Z_*P)$$

$$= \int_B \int_A p(y|z) dx d(Z_*P) = \int_B \int_A p(y|z) p(z) dz.$$

And therefore we a.s. have

$$p(y,z) = \frac{p(y|z)}{p(z)}.$$