Probabilities and Statistics I - Test II: Formulae

Random Variables and Joint Probability Distributions

Joint probability distribution/mass function (PDF/PMF): A function whose value at any given sample (X, Y) in the sample space can be interpreted as providing a relative likelihood that the value of the random variable would be equal to that sample.

1. $f(x,y) \geq 0 \; \forall \, x, y$
2. $\sum_x \sum_y f(x,y) = 1$ or $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dydx \, f(x) = 1$
3. $P(X = x, Y = y) = f(x,y)$ or $P(X \in [a,b], Y \in [c,d]) = \int_c^d \int_a^b dydx \, f(x,y)$

Marginal distributions of X alone given a joint probability distribution of X and Y:

$g(x) = \sum_y f(x,y), \; h(y) = \sum_x f(x,y)$ or $g(x) = \int_{-\infty}^{\infty} dy \, f(x,y), \; h(y) = \int_{\infty}^{\infty} dx \, f(x,y)$

X, Y are statistically independent $\iff f(x,y) = g(x)h(y)$.

Conditional distribution:

$$f(y|x) = \frac{f(x,y)}{g(x)}; \; g(x) > 0 \text{ and } f(x|y) = \frac{f(x,y)}{h(y)}; \; h(y) > 0$$

Mathematical Expectations

Mean Values (μ) or Expectations ($\mathbb{E}$):

$$\mu_x = \mathbb{E}[X] = \sum_x x f(x) \text{ or } \mu_x = \mathbb{E}[X] = \int_{-\infty}^{\infty} dx \, x f(x)$$

$$\textbf{Thm 4.1: } \mathbb{E}[g(X)] = \sum_x g(x) f(x) \implies \textbf{Thm 4.5: } \mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

$$\textbf{Thm 4.6: } \mathbb{E}[g(X) \pm h(X)] = \mathbb{E}[g(X)] \pm \mathbb{E}[h(X)]$$

$$\mathbb{E}[g(X,Y)] = \sum_x \sum_y g(x,y) f(x,y) \text{ or } \mu_{g(X,Y)} = \int_{-\infty}^{\infty} dxdy \, g(x,y) f(x,y)$$

$$\implies \textbf{Thm 4.7: } \mathbb{E}[g(X,Y) \pm h(X,Y)] = \mathbb{E}[g(X,Y)] \pm \mathbb{E}[h(X,Y)]$$

$$\implies \mathbb{E}[g(X) \pm h(Y)] = \mathbb{E}[g(X)] \pm \mathbb{E}[h(Y)]$$

$$\mathbb{E}[X,Y] = \mathbb{E}[X]\mathbb{E}[Y] \text{ if X, Y are independent}$$

Variance ($\sigma^2 = \mathbb{E}[(X - \mu)^2]$) is the expectation of the squared deviance $(X - \mu)^2$ and Standard Deviation ($\sigma = +\sqrt{\sigma^2}$) is the positive square-root of the variance:

$$\text{Var}(X) = \sigma^2 = \mathbb{E}[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x) \text{ or } \int_{-\infty}^{\infty} dx \, (x - \mu)^2 f(x)$$

$$\textbf{Thm 4.2: } \quad \text{Var}(X) = \sigma^2 = \mathbb{E}[X^2] - \mu_x^2$$

$$\textbf{Thm 4.3:}\quad \mathrm{Var}(g(X)) = \sigma^2_{g(X)} = \mathbb{E}[(g(X) - \mu_{g(X)})^2]$$

$$\therefore \mathrm{Var}(g(X)) = \sum_x [g(x) - \mu_{g(X)}]^2 f(x) \text{ or } \int_{-\infty}^{\infty} dx\, [g(x) - \mu_{g(X)}]^2 f(x)$$

Covariance ($\sigma_{X,Y}$) or ($\mathrm{Cov}(X,Y)$) is the expectation value of the joint variability of two random variables $\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$:

$$\sigma_{X,Y} = \sum_x \sum_y (x - \mu_X)(y - \mu_y) f(x,y) \text{ or } \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dxdy\, (x - \mu_X)(y - \mu_y) f(x,y)$$

$$\textbf{Thm 4.4:}\ \sigma_{X,Y} = \mathbb{E}[XY] - \mu_X \mu_Y$$

$$\textbf{Thm 4.9:}\quad \sigma^2_{aX \pm bY + c}. = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{X,Y} \implies \mathrm{Var}(aX + b) = a^2\sigma^2_X$$

The correlation coefficient ($\rho_{X,Y}$) is a normalized/scale-free measure of covariance:

$$\rho_{X,Y} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y};\quad \rho_{X,Y} = 0\text{: no linear relationship;}\quad \rho_{X,Y} = \pm 1\text{: exact linear relationship}$$

$$\textbf{Chebyshev's Theorem:}\quad P(X \in [\mu \pm k\sigma]) \geq 1 - \frac{1}{k^2} \iff P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Discrete Random Variables

Bernoulli processes and binomial distribution:

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x};\quad B(r; n, p) = \sum_{x=0}^{r} b(x; n, p);\quad \mu = np,\ \sigma = \sqrt{npq}$$

Hypergeometric distribution (x successes from n samples taken from k successes in N population):

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}};\quad \mu = \frac{nk}{N},\ \sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N}\left(1 - \frac{k}{N}\right)$$

Negative binomial/Pascal (k^{th} success in x Bernoulli trials) and geometric distributions:

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k};\quad b^*(x; 1, p) = g(x, p) = pq^{x-1};$$

$$\mu_{b^*} = \frac{k}{p},\ \sigma^2_{b^*} = \frac{k(1-p)}{p^2};\quad \mu_g = \frac{1}{p},\ \sigma^2_g = \frac{1-p}{p^2}$$

Poisson processes, distribution (x number of evenets occuring in λt time) and approximation to binomial distribution:

$$p(x; \lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!};\quad P(r; \lambda t) = \sum_{x=0}^{r} p(x; \lambda t);\quad \mu = \sigma^2 = \lambda t$$

$$\lim_{n\to\infty} \lim_{p\to 0} b(x; n, p) \to p(x; \lambda t);\quad \lambda t = np = \mu \text{ (constant)}$$