

# Named Probability Distributions

## Discrete Probability Distributions

pmf:  $p(y)$  cdf:  $F(y) = \sum_{z=-\infty}^y p(z)$   
 $0 \leq p(y) \leq 1; \sum_{y=-\infty}^{\infty} p(y) = 1$   
 $P(Y = y) = p(y); P(a \leq Y \leq b) = \sum_a^b p(y)$

### Binomial – $Y \sim \text{Binom}(n, p)$

$$p(y) = \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}, y \in [0, n], p \in [0, 1]$$
$$\mathbb{E}[Y] = np$$
$$\mathbb{V}[Y] = np(1-p)$$
$$m(t) = [pe^t + (1-p)]^n$$

### Geometric – $Y \sim \text{Geom}(p)$

$$p(y) = (1-p)^{y-1} p, y \in [1, \infty), p \in [0, 1]$$
$$\mathbb{E}[Y] = 1/p$$
$$\mathbb{V}[Y] = (1-p)/p^2$$
$$m(t) = \frac{p}{1-qe^t}$$

### Hypergeometric – $Y \sim \text{HG}(N, K, n)$

$$p(y = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}, k \in \{\max(0, n+K-N), \dots, \min(n, K)\}, K \leq N; n \leq N$$
$$\mathbb{E}[Y] = \frac{nK}{N}$$
$$\mathbb{V}[Y] = \frac{nK(N-K)(N-n)}{N^2(N-1)}$$

### Negative Binomial – $Y \sim \text{NBinom}(r, p)$

$$P(Y = k) = \binom{r+k-1}{r-1} p^r (1-p)^k, k \in [r, \infty), r \in \mathbb{Z}^+, p \in [0, 1]$$
$$\mathbb{E}[Y] = rq/p$$
$$\mathbb{V}[Y] = rq/p^2$$
$$m(t) = \left(\frac{p}{1-qe^t}\right)^r \text{ for } qe^t < 1$$

### Poisson – $Y \sim \text{Poi}(\lambda)$

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda}, y \in [0, \infty);$$
$$\mathbb{E}[Y] = \mathbb{V}[Y] = \lambda$$
$$m(t) = e^{\lambda(e^t-1)}$$

## Continuous Probability Distributions

pdf:  $f(y) = \frac{d}{dy}(y)$  cdf:  $F(y) = \int_{-\infty}^y f(z) dz$   
 $f(y) \geq 0; \int_{-\infty}^{\infty} f(y) dy = 1; P(Y = y) = 0$   
 $P(a \leq Y \leq b) = \int_a^b f(y) dy = F(b) - F(a)$

### Uniform – $Y \sim \text{Uniform}(a, b)$

$$f(y) = (b-a)^{-1}, y \in [a, b]$$
$$\mathbb{E}[Y] = (a+b)/2$$
$$\mathbb{V}[Y] = (b-a)^2/12$$
$$m(t) = (e^{bt} - e^{at})/[t(b-a)]$$

### Normal – $Y \sim N(\mu, \sigma^2)$

$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2} \quad y \in (-\infty, \infty), \mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$$
$$\mathbb{E}[Y] = \mu;$$
$$\mathbb{V}[Y] = \sigma^2$$
$$m(t) = \exp(\mu t + t^2 \sigma^2/2)$$

If  $Y \sim N(\mu, \sigma)$ , then  $Z = (Y - \mu)/\sigma; Z \sim N(0,1)$ .

$$P(Y \leq y) = \Phi\left(\frac{y-\mu}{\sigma}\right) = \Phi(z) \text{ (non-analytic function)}$$

### Exponential – $Y \sim \text{Exponential}(\lambda)$

$$f(y) = \lambda e^{-\lambda y}, y \in [0, \infty), \lambda \in \mathbb{R}^+$$
$$\mathbb{E}[Y] = 1/\lambda$$
$$\mathbb{V}[Y] = 1/\lambda^2$$
$$m(t) = \frac{\lambda}{\lambda - t} \text{ for } t < \lambda$$

### Gamma – $Y \sim \text{Gamma}(\alpha, \beta)$

$$f(y) = y^{\alpha-1} e^{-y/\beta} / [\beta^\alpha \Gamma(\alpha)], y \in [0, \infty), \alpha \in \mathbb{R}^+, \beta \in \mathbb{R}^+$$
$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy = (\alpha-1)\Gamma(\alpha-1)$$

If  $n$  is a positive integer,  $\Gamma(n) = (n-1)!$

$$\mathbb{E}[Y] = \alpha\beta$$
$$\mathbb{V}[Y] = \alpha\beta^2$$
$$m(t) = (1 - \beta t)^{-\alpha}$$

$\alpha = 1 \Rightarrow$  **exponential distribution**

$\beta = 2, \alpha = \nu/2, \nu \in \mathbb{Z}^+ \Rightarrow$  **chi-square distribution**

### Beta – $Y \sim \text{Beta}(\alpha, \beta)$

$$f(y) = y^{\alpha-1} (1-y)^{\beta-1} / B(\alpha, \beta), y \in [0, 1], \alpha \in \mathbb{R}^+, \beta \in \mathbb{R}^+$$
$$B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$$
$$\mathbb{E}[Y] = \alpha/(\alpha+\beta)$$
$$\mathbb{V}[Y] = \alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$$

### T – $Y \sim T(\nu)$

$$f(y) = \frac{\Gamma(\frac{\nu+1}{2})}{(\sqrt{\nu\pi}\Gamma(\nu/2))} \left(1 + \frac{y^2}{\nu}\right)^{-(\nu+1)/2}, y \in (-\infty, \infty), \nu > 0$$
$$\mathbb{E}[Y] = 0 (\nu > 1)$$
$$\mathbb{V}[Y] = \frac{\nu}{\nu-2} (\nu > 2), \infty (1 < \nu \leq 2)$$

# Properties of Estimators

## Inequalities and Convergence

**Cauchy-Schwarz:**  $|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$

**Jensen:**  $E(g(X)) \geq g(E(X))$  if  $g$  convex;  $E(g(X)) \leq g(E(X))$  if  $g$  concave.

**Markov:**  $P(|W| > a) \leq E(|W|)/a$

**Chebyshev:**  $P(|W - \mu| \geq \epsilon) \leq \sigma^2/\epsilon^2$

**Chernoff:**  $P(W \geq a) \leq E(e^{tW})/e^{ta}$

$X_n \rightarrow_d X$  if  $\lim_{n \rightarrow \infty} F_n(x) = F(x)$  at all  $x$

$X_n \rightarrow_p X$  if  $\lim_{n \rightarrow \infty} P(|X_n - X| \geq \epsilon) = 0$

## Fisher Information

$$I(\theta) = E\left[\left(\frac{\partial \ln f_Y(y; \theta)}{\partial \theta}\right)^2\right] = -E\left[\frac{\partial^2 \ln f_Y(y; \theta)}{\partial \theta^2}\right]$$

## Cramer-Rao Lower Bound

$Y_1, \dots, Y_n \sim f_Y, \{y : y \neq 0\}$  does not depend on  $\theta, E(\hat{\theta}) = \theta$ .  
 $Var(\hat{\theta}) \geq \frac{1}{nI(\theta)}$

## Consistency

$\hat{\theta}_n$  is consistent if  $\hat{\theta}_n \rightarrow_p \theta$ .

Invariance: If  $\hat{\theta}_n$  is consistent for  $\theta, g(\hat{\theta}_n)$  is consistent for  $g(\theta)$

## Sufficiency

$T = h(X_1, \dots, X_n)$  is sufficient for  $\theta$  if  $P(X_1, \dots, X_n | T = t)$  does not depend on  $\theta$ .

$T$  is sufficient if and only if  $L(\theta) = g[h(X_1, \dots, X_n); \theta] \cdot b(X_1, \dots, X_n)$

**Rao-Blackwell:** Let  $\hat{\theta}$  be an estimator of  $\theta$  with  $E(\hat{\theta}^2) < \infty$  and let  $T$  be a sufficient statistic. If  $\theta^* = E(\hat{\theta} | T = t)$ , then  $MSE(\theta^*, \theta) \leq MSE(\hat{\theta}, \theta)$ . Strict inequality unless  $\hat{\theta} = f(T)$ .

## Exponential Families

$$f(x; \theta) = \exp[\eta(\theta)T(x) - A(\theta) + B(x)]$$

$$= h(x) \exp[\eta(\theta)T(x) - A(\theta)]$$

$$= h(x)g(\theta) \exp[\eta(\theta)T(x)]$$

$$E(Y) = \frac{\partial}{\partial \eta} A(\eta)$$

$$V(Y) = \frac{\partial^2}{\partial \eta^2} A(\eta)$$

## Large-Sample Properties

**WLLN:**  $\bar{X} \rightarrow_p \mu$

**CLT:** If  $Y_i \sim f_Y, E(Y) = \mu, V(Y) = \sigma^2$ , then  $\bar{Y} \sim N(\mu, \sigma^2/n)$

**CLT:**  $\frac{\sum X_i - n\mu}{\sigma/\sqrt{n}} \rightarrow_d Z$ , where  $Z \sim N(0, 1)$ .

**Delta Method:** If  $Y_n \approx N(\mu, \frac{\sigma^2}{n})$  then  $g(Y_n) \approx N(g(\mu), (g'(\mu))^2 \frac{\sigma^2}{n})$

$\hat{\theta}_{MLE} \sim N(\theta, \frac{1}{nI(\theta)})$  for large  $n$

# Inference

## Sampling Distributions

If  $Z \sim N(0, 1), Z^2 \sim \chi^2(1)$

$\sum_{i=1}^n Z_i^2 \sim \chi^2(n)$

$[(n-1)S^2/\sigma^2] \sim \chi^2(n-1)$

$[(\bar{Y} - \mu)/(\sigma/\sqrt{n})] \sim N(0, 1)$

$[(\bar{Y} - \mu)/(S/\sqrt{n})] \sim t(n-1)$

$Z \sim N(0, 1) \perp W \sim \chi^2(\nu)$ , then  $T = (Z/\sqrt{W/\nu}) \sim t(\nu)$

$$S_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}, T = \frac{\bar{X} - \bar{Y} - \mu_x + \mu_y}{S_p \sqrt{1/n + 1/m}} \sim T_{n+m-2}$$

$$\frac{\bar{X} - \bar{Y} - \mu_x + \mu_y}{\sqrt{s_x^2/n + s_y^2/m}} \sim T[\nu], \nu = \frac{(s_x^2/s_y^2 + n/m)^2}{1/(n-1)(s_x^2/s_y^2 + 1/(m-1)(n/m)^2)}$$

$X \sim \text{Binom}(n, p_x), Y \sim \text{Binom}(m, p_y)$ ,

$X/n - Y/m \sim N(p_x - p_y, p_x(1-p_x)/n + p_y(1-p_y)/m)$

Goodness of Fit:  $\sum_{i=1}^k \frac{(X_i - n\hat{p}_i)^2}{n\hat{p}_i} \sim \chi_{k-1-s}^2$

## Power Function

Let  $\delta$  be a test of statistic  $T$  and rejection region  $R$ .

$$\pi(\theta|\delta) = P(T \in R|\theta)$$

## Likelihood Ratio Test (GLRT)

$$\lambda = \frac{\max_{\Omega_0} L(\theta)}{\max_{\Omega_1} L(\theta)}$$

$\delta = \text{Reject } H_0 \text{ when } \lambda \leq \lambda^*, \text{ where } P(\Lambda \leq \lambda^* | \theta \in \theta_0) = \alpha$

# Regression

## Simple Linear Regression

$$\hat{\beta}_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 \sim N(\beta_1, \frac{\sigma^2}{\sum (x_i - \bar{x})^2})$$

$$\hat{\beta}_0 \sim N(\beta_0, \sigma^2 [\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}])$$

$$\frac{S^2}{\sigma^2} \sim \chi_{n-2}^2, \hat{y}_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2 [\frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}])$$

$$\hat{y}_i^* - y^* \sim N(0, \sigma^2 [1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}])$$

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

$$r = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

$$R^2 = r^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

## Multiple Regression

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

$$\hat{\beta} \sim MVN(\beta, \sigma^2 (X^T X)^{-1})$$

$$\hat{\beta}_j \sim N(\beta_j, \sigma^2 c_{jj})$$

$$H = X(X^T X)^{-1} X^T Y$$

$$H = H^T = H^2 \text{ and } (I - H) = (I - H)^T = (I - H)^2$$

$$\frac{S^2}{\sigma^2} \sim \chi_{n-p}^2$$

$$\Sigma_{\hat{\epsilon}} = \sigma^2 (I - H)$$

$$\text{Adj } R^2 = 1 - \frac{(1/(n-p)) \sum (y_i - \hat{y}_i)^2}{(1/(n-1)) \sum (y_i - \bar{y})^2}$$

---

### Mean and Variance of Vector RVs

If  $E(\mathbf{Y}) = \boldsymbol{\mu}$ ,  $\text{Cov}(\mathbf{Y}) = \boldsymbol{\Sigma}_Y$ , and  $\mathbf{Z} = \mathbf{c} + \mathbf{A}\mathbf{Y}$ , then  
 $E(\mathbf{Z}) = \mathbf{c} + \mathbf{A}E(\mathbf{Y})$  and  $\boldsymbol{\Sigma}_Z = \mathbf{A}\boldsymbol{\Sigma}_Y\mathbf{A}^T$