MAP 6111 CLASS NOTES 8/23

REVIEW OF PROBABILITY

PROBABILITY SPACE: $(\Omega, \mathcal{F}, P)$

$\Omega = \{$SPACE OF OUTCOMES$\}$, EXAMPLE: $\Omega = \{H, T\}$

$\mathcal{F}$ = THE SET OF SUBSETS OF Ω TO WHICH WE ASSIGN PROBABILITIES.

EXAMPLE: $\mathcal{F} = \{\phi, \{H\}, \{T\}, \{H \cup T\}\}$ ELEMENTS OF $\mathcal{F}$ ARE CALLED "EVENTS"

$P(H) = 1/2$, $P(T) = 1/2$, $P(\emptyset) = 0$, $P(H \cup T) = 1$

ϕ IS CALLED THE "IMPOSSIBLE EVENT"

Ω IS THE "SURE EVENT"

SUPPOSE $\Omega = \{[0,1]\}$, $\mathcal{F} = \{$ALL POSSIBLE INTERVALS, THEIR UNIONS AND INTERSECTIONS$\}$

$P([a,b]) = b - a$ OR $P([a,b]) = F(b) - F(a)$

$F(x)$ IS A CONTINUOUS FUNCTION.

IF $A \in \mathcal{F}$, A IS AN EVENT, THEN ① $0 \leq P(A) \leq 1$

IF $A \cap B = \emptyset$, THEN $P(A \cup B) = P(A) + P(B)$ ②

$P(\emptyset) = 0$, $P(\Omega) = 1$ ③

EVENTS A AND B ARE INDEPENDENT IF $P(A \cap B) = P(A) \cdot P(B)$.

EVENTS A, B, AND C ARE INDEPENDENT IF THEY ARE PAIRWISE INDEPENDENT AND

$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$.

CONDITIONAL PROBABILITY

$P(A|B)$ IS THE CONDITIONAL PROBABILITY OF EVENT A GIVEN EVENT B (AND $P(B) \neq 0$).

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{OR} \quad P(B|A) = \frac{P(A \cap B)}{P(A)} \rightarrow P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

MAP 6111 CLASS NOTES 8/25

REVIEW OF PROBABILITY (CONTINUED)

CHARACTERISTICS OF RANDOM VARIABLES

EXPECTATION: $E[X]$

IF X IS DISCRETE WITH $P(X=x_i)=p_i$, $i=1,2,\ldots$ THEN $E[X]=\sum_{i=1}^{\infty} x_i p_i$

IF X IS CONTINUOUS WITH PROBABILITY DENSITY FUNCTION $p(x)$, THEN
$$E[X]=\int_{-\infty}^{\infty} x\,p(x)\,dx.$$

PROPERTIES $E[aX+b] = aE[X]+b$, WHERE a, b ARE NUMBERS

FOR TWO RANDOM VARIABLES X_1 AND X_2,

$$E[X_1+X_2]=E[X_1]+E[X_2]$$

VARIANCE

$$VAR(X) = E[X-E[X]]^2$$

FOR X DISCRETE, $VAR(X)=\sum_{i=1}^{\infty}(x_i-E[X])^2 p_i$

FOR X CONTINUOUS, $VAR(X)=\int_{-\infty}^{\infty}(x-E[X])^2 p(x)\,dx$

EXPECTATION IS A MEASURE OF LOCATION OF X.

VARIANCE IS A MEASURE OF VARIABILITY

PROPERTIES OF VARIANCE

$$VAR(aX+b) = a^2\,VAR(X)$$

IF X_1 AND X_2 ARE INDEPENDENT, THEN $VAR(X_1+X_2) = VAR(X_1)+VAR(X_2)$

COVARIANCE

$$COV(X_1,X_2) = E[(X_1-E[X_1])(X_2-E[X_2])]$$

IF X IS CONTINUOUS, THEN

$$\varphi_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} p(x)\,dx \rightarrow \text{THE FOURIER TRANSFORM OF } p(x)$$

$$\varphi_X'(\omega) = \int_{-\infty}^{\infty} i x p(x) e^{i\omega x}\,dx$$

$$E[X] = \frac{1}{i}\varphi_X'(\omega)\Big|_{\omega=0}$$

$$E[X^k] = \frac{1}{i^k}\cdot\frac{d^k}{d\omega^k}\varphi_X(\omega)\Big|_{\omega=0}$$

DISCRETE DISTRIBUTIONS

DISCRETE UNIFORM DISTRIBUTION

X HAS A DISCRETE UNIFORM DISTRIBUTION $[1,\dots,N]$ IF $P(X=x/N) = 1/N$, $x=1,\dots,N$

HYPERGEOMETRIC DISTRIBUTION

SUPPOSE THERE IS AN URN WITH N BALLS, M OF THEM ARE RED AND (N−M) ARE WHITE. WE SELECT K BALLS BLINDFOLDED. WHAT IS THE PROBABILITY THAT EXACTLY x BALLS OUT OF K ARE RED?

$$P(X=x \mid N,M,K) = \frac{\text{NUMBER OF SAMPLES WITH } x \text{ RED BALLS}}{\text{TOTAL NUMBER OF SAMPLES}}$$

$$= \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} \qquad x = 0,\dots,K$$

$$\max(0, K+M-N) \le x \le \min(K, M)$$

TO FIND $E[X]$, USE THE FOLLOWING RELATION

$$\binom{L}{K} = \frac{L}{K}\cdot\binom{L-1}{K-1}$$

ξ	0	1
$P(\cdot)$	q	p

$$E[\xi] = 0 \cdot q + 1 \cdot p = p$$

$$\begin{aligned} \mathrm{VAR}(\xi) = E[\xi - E[\xi]]^2 &= E\left[\xi^2 - 2\xi E[\xi] + (E[\xi])^2\right] \\ &= E[\xi^2] - 2E[\xi]E[\xi] + (E[\xi])^2 \\ &= E[\xi^2] - (E[\xi])^2 \end{aligned}$$

TRUE FOR ANY RANDOM VARIABLE

$$\begin{aligned} &= p - p^2 \\ &= pq \end{aligned}$$

$$X = \sum_{i=1}^{n} \xi_i \quad \rightarrow \quad E[X] = E\left[\sum_{i=1}^{n} \xi_i\right] = np$$

$$\mathrm{VAR}(X) = npq \quad (\text{INDEPENDENCE } \mathrm{VAR}(\Sigma) = \Sigma\, \mathrm{VAR})$$

POISSON DISTRIBUTION

CONSIDER AN EXPERIMENT WHERE YOU ARE WAITING FOR AN OCCURENCE OF SOME EVENT.

ASSUMPTIONS:

① WITHIN A SMALL INTERVAL OF TIME, THE PROBABILITY THAT EVENT OCCURS IS PROPORTIONAL TO THE LENGTH OF TIME INTERVAL

② TWO EVENTS CANNOT OCCUR SIMULTANEOUSLY

$X = \{$ THE NUMBER OF OCCURENCES WITHIN A FIXED TIME INTERVAL $\}$

X HAS POISSON DISTRIBUTION:

$$P(X=x/\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x=0,1,2,\ldots \qquad \underbrace{\lambda}_{\text{POISSON INTENSITY}} = \lambda_0 t_0$$

FIND CHARACTERISTIC FUNCTION:

$$\varphi_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx} \cdot \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} \frac{e^{-\lambda}(\lambda e^{it})^x}{x!}$$

$$= e^{-\lambda} e^{\lambda e^{it}} \underbrace{\sum_{x=0}^{\infty} \frac{(\lambda e^{it})^x}{x!} \cdot e^{-(\lambda e^{it})}}_{\sum_{x=0}^{\infty} P(X=x/\lambda e^{it}) = 1}$$

$$= e^{-\lambda(1-e^{it})}$$

$$= e^{\lambda(e^{it}-1)}$$

$$\varphi_X(t) = e^{\lambda(e^{it}-1)}$$

$$\varphi_X'(t) = e^{\lambda(e^{it}-1)} \cdot \lambda i e^{it}$$

$$\varphi_X''(t) = e^{\lambda(e^{it}-1)}(\lambda i e^{it})^2 + e^{\lambda(e^{it}-1)}\lambda i^2 e^{it}$$

$$E[X] = \frac{1}{i}\varphi_X'(0) = \frac{\lambda i}{i} = \lambda$$

$$E[X^2] = \frac{1}{i^2}\varphi_X''(0) = \lambda^2+\lambda$$

$$VAR(X) = E[X^2] - [E[X]]^2 = \lambda^2+\lambda-\lambda^2 = \lambda$$

NEGATIVE BINOMIAL DISTRIBUTION

CONSIDER THE SEQUENCE OF BERNOULLI TRIALS (INDEPENDENT, WITH PROBABILITY p OF SUCCESS IN EACH TRIAL).

$X = \{$THE NUMBER OF A TRIAL WHERE THE r^{th} SUCCESS IS OBTAINED$\}$

$$P(X=x \mid r,p) = \binom{x-1}{r-1} p^{r-1}(1-p)^{(x-1)-(r-1)} \cdot p \,, \quad x = r, r+1, r+2, \ldots$$

$(x-1)$ TRIALS WITH $(r-1)$ SUCCESSES + LAST SUCCESS

$$P(X=x \mid r,p) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \,, \quad x = r, r+1, \ldots$$

IF $Y = x-r$ = THE NUMBER OF FAILURES BEFORE THE SUCCESS OCCURS,

THEN $P(Y=y \mid r,p) = \binom{y+r-1}{r-1} p^r (1-p)^y \,, \quad y = 0,1,\ldots$

$$E[Y] = \frac{r(1-p)}{p} \,, \quad VAR(Y) = \frac{r(1-p)}{p^2}$$

SINCE $X = Y+r$

$$E[X] = \frac{r(1-p)}{p} + r \,, \quad VAR(X) = VAR(Y)$$

CONTINUOUS DISTRIBUTIONS

① UNIFORM DISTRIBUTION

$$f(x|a,b) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0 & \text{OTHERWISE} \end{cases}$$

$$E[X] = \int_a^b x f(x|a,b)\,dx = \frac{1}{b-a}\int_a^b x\,dx = \frac{b+a}{2}$$

$$VAR(X) = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{1}{b-a}\,dx = \frac{(b-a)^2}{12}$$

$$\varphi_X(t) = \int_a^b e^{itx} \cdot \frac{1}{b-a}\,dx = \frac{1}{it(b-a)} \cdot \left(e^{itb} - e^{ita}\right)$$

② GAMMA DISTRIBUTION

RELIABILITY, SURVIVAL ANALYSIS

RECALL GAMMA FUNCTION:

$$\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1}\,dx, \quad \alpha > 0$$

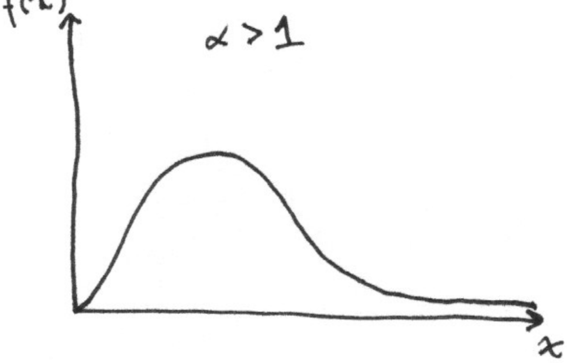
IF $\alpha = n \in \mathbb{Z}$, $\Gamma(n) = (n-1)!$

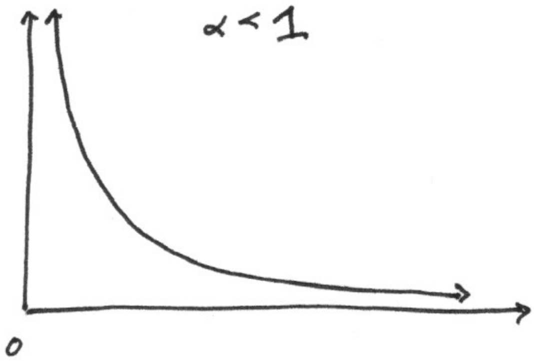
FOR ANY $\alpha > 0$, $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

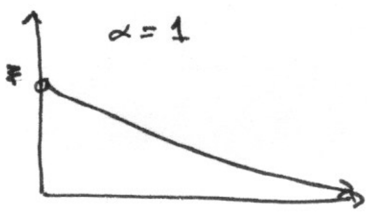
THE FAMILY OF GAMMA DISTRIBUTIONS WITH PARAMETERS α AND β HAS THE PROBABILITY DENSITY FUNCTION (pdf)

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)} \cdot \frac{x^{\alpha-1}}{\beta^\alpha} \cdot e^{-\frac{x}{\beta}} \quad ; \; x \geq 0, \alpha > 0, \beta > 0$$

~~LET~~ SUPPOSE X = A LIFETIME OF A CERTAIN DEVICE, ~~$\alpha > 1$~~







MAP6111 CLASS NOTES 9/1/2010

THE GAMMA DISTRIBUTION

USED IN RELIABILITY, SURVIVAL ANALYSIS

THE GAMMA FUNCTION

$$\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1}\, dx, \quad \alpha > 0$$

IF α IS AN INTEGER, THEN $\Gamma(\alpha) = (\alpha-1)!$.

FOR ANY $\alpha > 0$, $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.

THE FAMILY OF GAMMA DISTRIBUTIONS WITH PARAMETERS α AND β HAS THE P.D.F.

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)} \cdot \frac{x^{\alpha-1}}{\beta^\alpha} \cdot e^{-\frac{x}{\beta}}, \quad x \geq 0;\ \alpha > 0,\ \beta > 0$$

α IS THE "SHAPE" PARAMETER

β IS THE "SCALE" ~~"SKEW"~~ PARAMETER

$$E[X^k] = \int_0^\infty x^k \cdot \frac{1}{\Gamma(\alpha)} \cdot \frac{x^{\alpha-1}}{\beta^\alpha} \cdot e^{-x/\beta}\, dx$$

$$= \int_0^\infty \frac{x^{k+\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \cdot e^{-x/\beta}\, dx$$

LET $z = x/\beta \leftrightarrow x = \beta z$

$dx = \beta\, dz$

$$= \int_0^\infty \frac{\beta^{k+\alpha-1}}{\Gamma(\alpha)\beta^\alpha} \cdot e^{-z} \cdot z^{k+\alpha-1} \cdot \beta\, dz$$

$$= \frac{\beta^k}{\Gamma(\alpha)} \underbrace{\int_0^\infty e^{-z} z^{k+\alpha-1}\, dz}_{\Gamma(k+\alpha)}$$

$$= \frac{\beta^k \Gamma(k+\alpha)}{\Gamma(\alpha)}$$

<u>$k=1$</u> (MEAN)

$$E[X] = \frac{\beta\Gamma(\alpha+1)}{\Gamma(\alpha)} = \beta\alpha$$

<u>$k=2$</u> (USED TO FIND VARIANCE)

$$E[X^2] = \frac{\beta^2\Gamma(\alpha+2)}{\Gamma(\alpha)} = \beta^2(\alpha^2+\alpha)$$

$$\Rightarrow VAR(X) = E[X^2] - (E[X])^2$$

$$= \beta^2(\alpha^2+\alpha) - \beta^2\alpha^2$$

$$= \alpha\beta^2$$

$$\phi(t) = E[e^{itX}] = \int_0^\infty \frac{e^{itx}}{\Gamma(\alpha)} \cdot \frac{x^{\alpha-1}}{\beta^\alpha} \cdot e^{-\frac{x}{\beta}}\, dx$$

$$\downarrow z = x\left(\tfrac{1}{\beta} - it\right)$$

$$\phi(t) = \frac{1}{(1-i\beta t)^\alpha}$$

GENERALIZATION: WEIBULL DISTRIBUTION

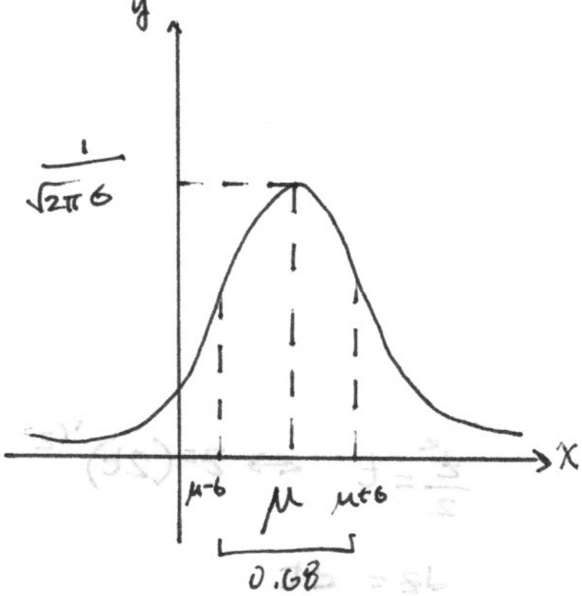
$X \sim \text{EXPONENTIAL}(\beta) \rightarrow Y = X^{\frac{1}{\gamma}} \sim \text{WEIBULL}(\gamma, \beta)$

CHI-SQUARED DISTRIBUTION

$\chi^2(p) = \text{GAMMA}(p/2, 2)$ $\quad \alpha = p/2$, $\beta = 2$, p IS DEGREES OF FREEDOM

NORMAL (OR GAUSSIAN) DISTRIBUTION

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$ IS THE P.D.F. OF $N(\mu, \sigma^2)$



$P(|X-\mu| \le \sigma) = 0.6826$

$P(|X-\mu| \le 2\sigma) = 0.9544$

$P(|X-\mu| \le 3\sigma) = 0.9974$

THE CENTRAL LIMIT THEOREM

IF $X_1, \dots, X_n$ ARE INDEPENDENT WITH $E[X_i] = \cancel{\mu_i}\ m_i$ AND $VAR(X_i) = \sigma_i^2$, THEN, UNDER SOME CONDITIONS,

$$Z_n = \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n m_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$, WE HAVE $$\lim_{n\to\infty} P(Z_n > z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} \cdot e^{-t/2}\, dt = P(Z > z)$$

Z IS THE STANDARD NORMAL VARIABLE, $Z \sim N(0,1)$

$$E[X] = \mu + \sigma E[Z] = \underset{\text{MEAN}}{\mu}\ ; \quad VAR(X) = \sigma^2 VAR(Z) = \overset{\text{VARIANCE}}{\sigma^2}$$

$$\Phi(x) = P(Z \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$ CUMULATIVE DISTRIBUTION FUNCTION

NORMAL APPROXIMATION OF BINOMIAL PROBABILITIES

$X \sim$ BINOMIAL (n, p)

$$P(X = x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

$X = \sum_{i=1}^{n} X_i$, X_i ARE INDEPENDENT BERNOULLI VARIABLES

$$E[X_i] = p, \quad VAR(X_i) = p \cdot q \quad ; \ p = P(X_i = 1)$$

$$\frac{X - np}{\sqrt{npq}} \rightarrow N(0,1) \longleftrightarrow P\left(\frac{X-np}{\sqrt{npq}} < z\right) \xrightarrow[n \to \infty]{} \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$

$$P(X = x) = P(x - 1/2 \leq X \leq x + 1/2)$$

$$= P\left(\underbrace{\frac{x - 1/2 - np}{\sqrt{npq}}}_{z_1} \leq \underbrace{\frac{X - np}{\sqrt{npq}}}_{Z} \leq \underbrace{\frac{x + 1/2 - np}{\sqrt{npq}}}_{z_2}\right)$$

$$= \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$

$$= \Phi(z_2) - \Phi(z_1)$$